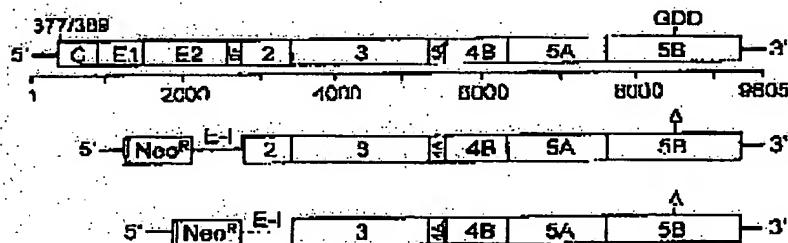


Français
StrategisContact Us
Site MapHelp
What's NewSearch
About UsCanada Site
Registration**Strategis.gc.ca****CIPO** **OPIC**
CIPO Home

Stratégis Index: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Canadian Intellectual Property Office

Canadian Patents Database**(12) Patent Application:****(11) CA 2303526****(54) HEPATITIS C VIRUS CELL CULTURE SYSTEM****(54) SYSTEME DE CULTURE DU VIRUS DE L'HEPATITE C****Representative Drawing:****View or Download Images****View Administrative Status****ABSTRACT:**

The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).

CLAIMS: Show all claims

*** Note: Data on abstracts and claims is shown in the official language in which it was submitted.

(72) Inventors (Country): **BARTENSCHLAGER, RALF** (Germany (Federal Republic of))

(73) Owners (Country): **BARTENSCHLAGER, RALF** (Germany (Federal Republic of))

(71) Applicants (Country): **BARTENSCHLAGER, RALF** (Germany (Federal Republic of))

(74) Agent: **BORDEN LADNER GERVAIS LLP**

(45) Issued:

(22) Filed: **Mar. 31, 2000**

(41) Open to Public Inspection: **Oct. 3, 2000**

Examination requested: **Jan. 11, 2002**

(51) International Class (IPC): **C12N 7/02
C12N 5/10
C12N 15/51**

Patent Cooperation Treaty (PCT): No

(30) Application priority data:

Application No.	Country	Date
199 15 178.4	Germany (Federal Republic of)	Apr. 3, 1999

Availability of licence: **N/A**

Language of filing: **English**

View or Download Images :

- Cover Page Image
- Abstract Image
- Claims Image
- Disclosures Image
- Drawings Image
- Representative Drawing Image

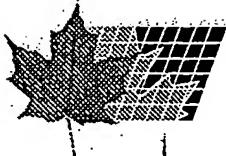
[View the image](#)

[Download in Adobe PDF](#)

MicroPatent® Family Lookup

Stage 1 Patent Family - "Complex"		Priorities and Applications	
CC DocNum	KD PubDate	CC AppNum	KD AppDate
<input type="checkbox"/> AT 236988	E 20030415	AT 105929	EP 20000323
		DE 19915178	A 19990403
<input type="checkbox"/> AU 200025180	A5 20001019	AU 25180	A 20000331
		DE 19915178	A 19990403
<input type="checkbox"/> CA 2303526	AA 20001003	CA 2303526	A 20000331
		DE 19915178	A 19990403
<input type="checkbox"/> DE 19915178	A1 20001005	DE 19915178	A 19990403
		DE 19915178	A 19990403
<input type="checkbox"/> DE 50001673	C0 20030515	DE 19915178	A 19990403
		DE 50001673	A 20000323
		DE 50001673	A 20000323
<input type="checkbox"/> DK 1043399	T3 20030728	DE 19915178	A 19990403
		DK 105929	A 20000323
<input type="checkbox"/> EP 1043399	A2 20001011	DE 19915178	A 19990403
		EP 105929	A 20000323
<input type="checkbox"/> EP 1043399	B1 20030409	DE 19915178	A 19990403
		EP 105929	A 20000323
<input type="checkbox"/> EP 1043399	A3 20020508	DE 19915178	A 19990403
		EP 105929	A 20000323
<input type="checkbox"/> JP 2001017187	A2 20010123	DE 19915178	A 19990403
		JP 2000101615	A 20000403

10 Publications found.



(21) (A1) 2,303,526

(22) 2000/03/31

(43) 2000/10/03

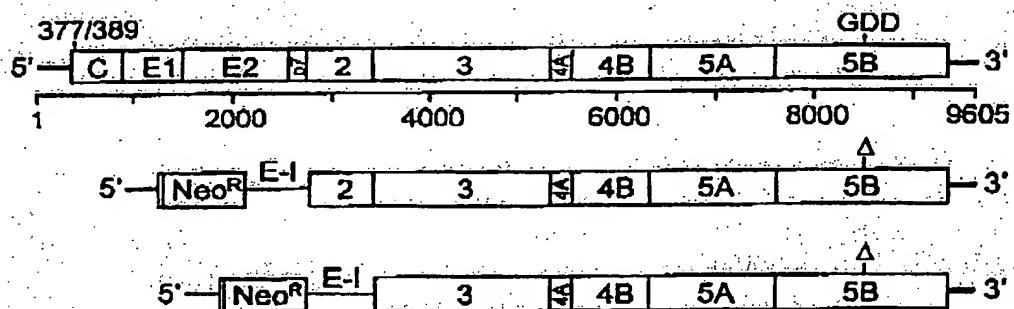
(72) BARTENSCHLAGER, RALF, DE
(71) BARTENSCHLAGER, RALF, DE

(51) Int.Cl. ⁷ C12N 7/02, C12N 15/51, C12N 5/10

(30) 1999/04/03 (199 15 178.4) DE

(54) SYSTÈME DE CULTURE DU VIRUS DE L'HEPATITE C

(54) HEPATITIS C VIRUS CELL CULTURE SYSTEM



(57) The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).



Hepatitis C Virus Cell Culture System

ABSTRACT

The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises 10 the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).

CLAIMS

1. A hepatitis C virus (HCV) cell culture system, which comprises mainly eukaryotic cells containing transfected HCV specific genetic material, characterized in that,
the eukaryotic cells are human hepatoma cells and the transfected HCV specific genetic material is a HCV-RNA construct, which comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as an additional marker gene for selection (selection gene).
2. A cell culture system according to Claim 1, characterized in that,
the hepatoma cells are derived from a usual hepatoma cell line.
3. A cell culture system according to Claim 1, characterized in that,
the hepatoma cells are extracted from a primary hepatoma cell culture.
4. A HCV-RNA construct, characterized in that,
it comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as an additional marker gene for selection (selection gene).
5. A HCV-RNA construct according to Claim 4, characterized in that,
it comprises one of the nucleotide sequences illustrated in the sequence protocols SEQ ID NO: 1 to SEQ ID NO: 11.
- 35 6. A HCV-RNA construct according to Claim 4, characterized in that,

the 3' NTR has a nucleotide sequence, which has been selected from the group of nucleotide sequences (a) to (i) listed in the following:

- 5 (a) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTA
GCTTTTTTTTTCTTTTTGAGAGAGAGTCTCACTCTGTTGCC
CAGACTGGAGT

10 (b) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTA
GTCTTTTTTTCTTTTTGAGAGAGAGTCTCACTCTGTTGCCA
GAATGGAGC

15 (c) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTAAT
CTTTTTTTTTCTTTTTGAGAGAGAGAGTCTCACTCTGTTGCCA
GAATGCAGC

20 (d) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
AGTCTTTTTTTCTTTTTGAGAGAGAGAGTCTCACTCTGTTGC
CCAGACTGGAGT

25 (e) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTA
GTCTTTTTTTCTTTTTGAGAGAGAGAGTCTCACTCTGTTGC
CCAGACTGGAGT

30 (f) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTA
GTCTTTTTTTCTTTTTGAGAGAGAGAGTCTCACTCTGTTG
CCCAGACTGGAGT

(g) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTA
GTCTTTTTTTCTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC
CAGACTGGAGT

35 (h) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
TAATCTTTTTTTCTTTTTGAGAGAGAGAGTCTCACTCTGT
TGCCCAGACTGGAGT

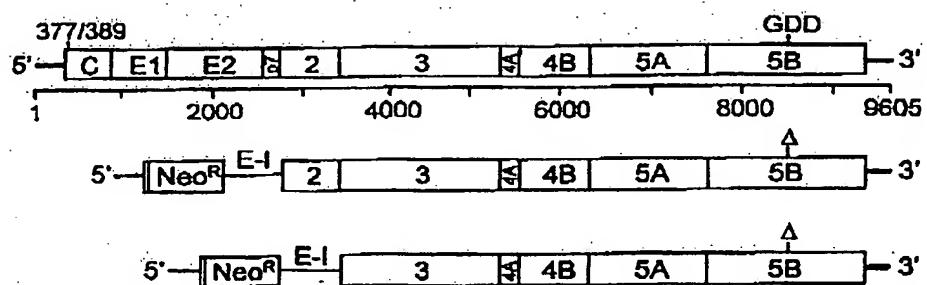
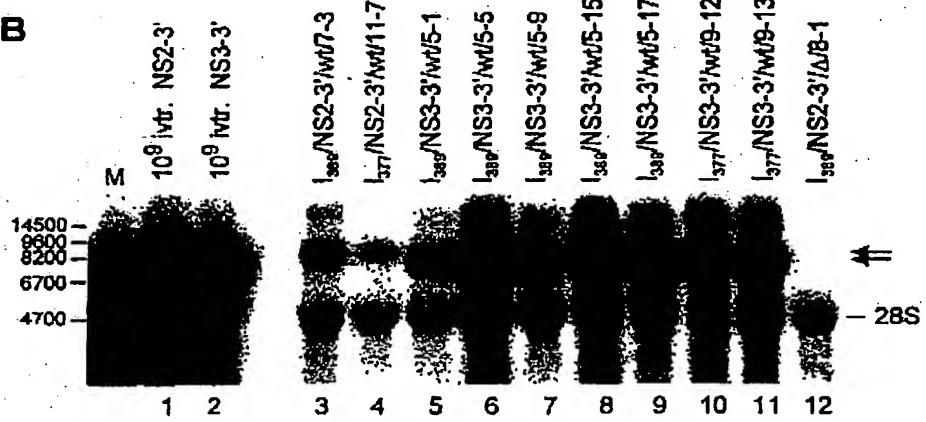
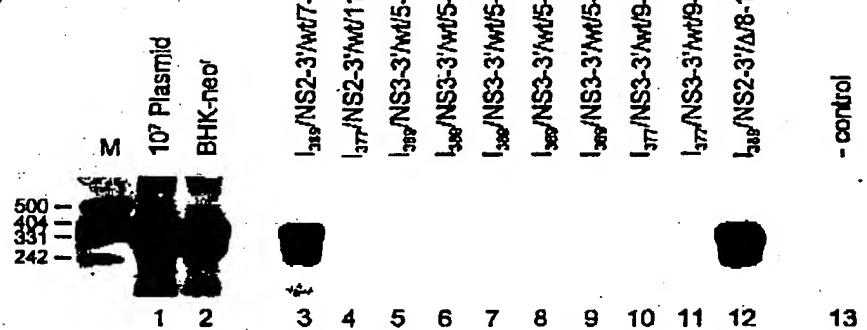
- (i) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
AATCTTTTTTTTTTTCTTTTTTGAGAGAGAGTCTCACTCTG
TTGCCAGACTGGAGT
- 5 7. A HCV-RNA construct according to one of the Claims 4 to 6, characterized in that,
the marker gene for selection is a resistance gene and preferably an antibiotic resistance gene.
- 10 8. A HCV-RNA construct according to one of the Claims 4 to 6, characterized in that,
the marker gene for selection is a neomycin phosphotransferase gene.
- 15 9. A HCV-RNA construct according to one of the Claims 4 to 8, characterized in that,
the marker gene for selection is integrated in the HCV-RNA downstream from the 5' NTR.
- 20 10. A HCV-RNA construct according to one of the Claims 4 to 83, characterized in that,
the marker gene for selection is linked to the HCV-RNA via a ribozyme or a recognition sequence for a ribozyme.
- 25 11. A HCV-RNA construct according to one of the Claims 4 to 10, characterized in that,
it has an integrated reporter gene.
- 30 12. A HCV-RNA construct according to Claim 11, characterized in that,
the reporter gene is a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) and the SEAP gene (secreted alkaline phosphatase gene).

13. A HCV-RNA construct according to one of the Claims 4 to 11, characterized in that, their replication has an influence on the expression of a (cellular) surrogate marker gene.
14. A HCV-RNA construct according to one of the Claims 11 to 13, characterized in that, the resistance gene is cloned into the open reading frame of the HCV-RNA in such a way that it will only be transferred to an active form after proteolytic processing.
15. A HCV-RNA construct according to one of the Claims 11 to 14, characterized in that, the reporter gene and the marker gene for selection are arranged in the construct in such a way, that they are expressed as a fusion protein.
16. A cell culture system according to one of the Claims 1 to 3, characterized in that, the HCV-RNA construct is a construct according to at least one of the Claims 4 to 15.
17. A cell culture system according to Claim 1, characterized in that, the cells containing the HCV-RNA construct are lodged at the DSMZ, Braunschweig, Germany under the deposit number DSM ACC2394 (laboratory name HuBl 9-13).
18. A use of a cell culture system according to one of the Claims 1 to 3 or 16 to 17 and/or a HCV-RNA construct according to one of the Claims 4 to 15 for the production and/or evaluation and/or testing of treatments and/or diagnostics for the treatment of HCV infections in particular.

19. A use of a cell culture system according to one of the Claims 1 to 3 or 16 to 17 and/or a HCV-RNA construct according to one of the Claims 4 to 15 for the production of a vaccine against HCV infections.
- 5
20. A use a HCV-RNA construct according to one of the Claims 4 to 15 for the production of a liver cell specific vector in gene therapy.
- 10 21. A HCV-RNA construct according to one of the Claims 4 to 15, characterized in that, it has an integrated foreign gene and can be used to transfect this foreign gene in a target cell, which can be used to express this foreign gene.
- 15
22. A process for the production of cell culture adapted mutants of a HCV-RNA construct according to one of the claims 4 to 15, in which the mutants have increased replication efficiency compared to the HCV-RNA construct,
characterized in that,
a cell culture system according to claim 1, in which the transfected HCV specific genetic material is a HCV-RNA construct with a selection gene according to one of the claims 4 to 15, is cultivated on/in the selection medium corresponding to the selection gene, that the cell clones created are collected, and that the HCV-RNA constructs or parts thereof are isolated from these cell clones.
- 20
25
30
35
23. A process according to claim 22, characterized in that,
the isolated HCV-RNA constructs are passaged at least one more time. That is, they are transfected in cells of a cell culture system according to claim 1. The cell culture system of claim 1 thereby obtained, in which the transfected HCV specific genetic material is

- the isolated HCV-RNA construct with a selection gene, is cultivated on/in the selection medium corresponding to the selection gene, the cell clones created are collected, and the HCV-RNA constructs are isolated from these cell clones.
24. A process for the production of mutants of a HCV full-length genome or of a HCV subgenome or any HCV construct with increased replication efficiency in comparison to the original HCV full-length genome or subgenome or HCV-RNA construct,
characterized in that,
a cell culture adapted mutant of a HCV-RNA construct is produced and isolated by a process according to claim 22 or 23,
the nucleotide and amino acid sequence of these mutants is determined and the type, number and positions of the nucleotide and amino acid mutations is determined by comparison with the nucleotide and amino acid sequence of the original HCV-RNA construct, and these mutations are introduced in an (isolated) HCV full-length genome or a HCV subgenome or any HCV-RNA construct, either by site-directed mutagenesis or by exchange of fragments containing the relevant mutations
25. A cell culture adapted HCV-RNA construct with high replication efficiency,
characterized in that,
it is derivable from a HCV-RNA construct according to one of the claims 4 to 15 through nucleotide and/or amino acid mutations and that it is obtainable with a process according to one of the claims 22 to 24.
- 30
35 26. A cell culture adapted HCV-RNA construct according to claim 25,
characterized in that,

- it carries one or several of the following amino acid exchanges: 1283 arg -> gly and/or 1383 glu -> ala and/or 1577 lys -> arg and/or 1609 lys -> glu and/or 1936 pro -> ser and/or 2163 glu -> gly and/or 2330 lys -> glu and/or 2442 ile -> val.
- 5
27. A cell culture adapted HCV-RNA construct according to claim 25 or 26,
characterized in that,
10 it displays one or several of the nucleotide and/or amino acid exchanges shown in Table 3, where Table 3 is an integral part of this claim.
- 15
28. Cell culture adapted mutant of a HCV-RNA construct or of a HCV full-length genome with increased replication efficiency in comparison to the original HCV-RNA construct or original HCV full-length genome, characterized in that,
20 it is obtainable with a process in which the number and type of mutations in a cell culture adapted HCV-RNA construct according to claim 24 are determined through sequence analysis and sequence comparison, and these mutations are introduced in a HCV-RNA construct, particularly in a HCV-RNA construct according to one of the claims 4 to 15, or in an (isolated) HCV-RNA full-length genome, either by targeted mutagenesis or by exchange of sequence sections containing the relevant mutations.
- 25
- 30 29. Hepatitis C virus particles or virus-like particles, characterized in that,
they are obtainable by a process according to one of the claims 22 - 24.
- 35 30. Cells infected with Hepatitis C virus particles or virus-like particles, according to claim 29.

A**B****C****Fig. 1**

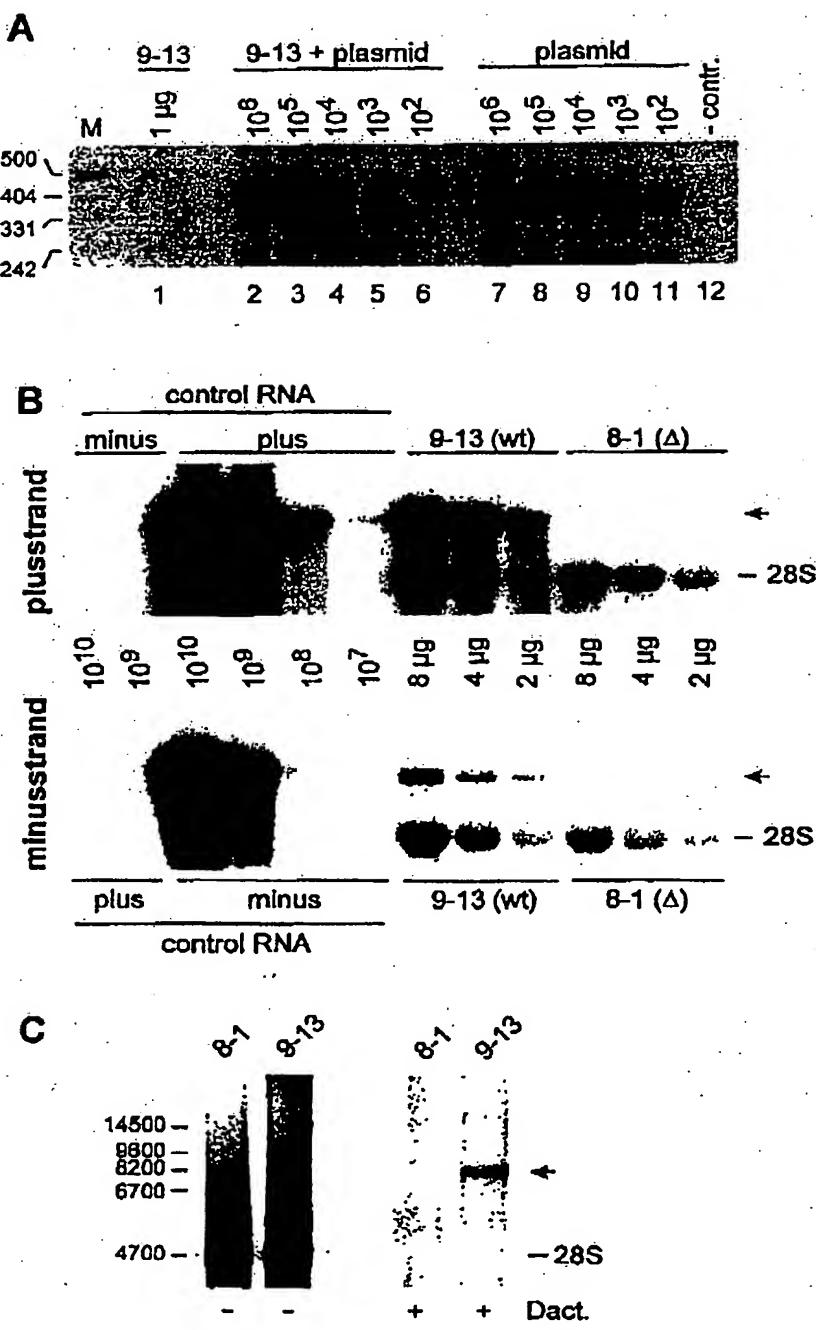


Fig. 2

3
Fig.

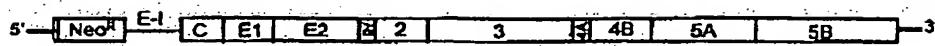
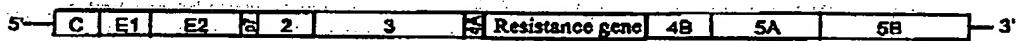
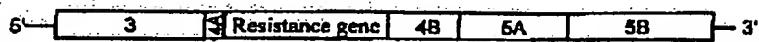


Fig. 4

A



B

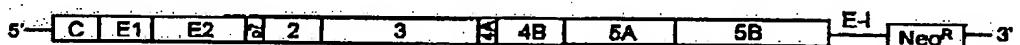
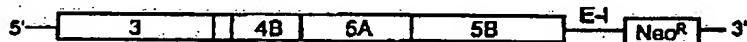
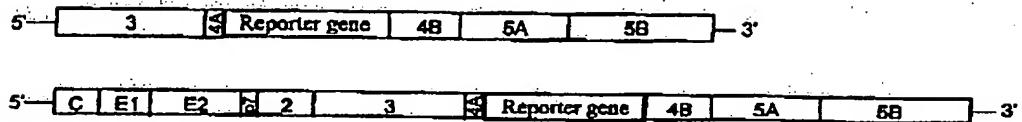
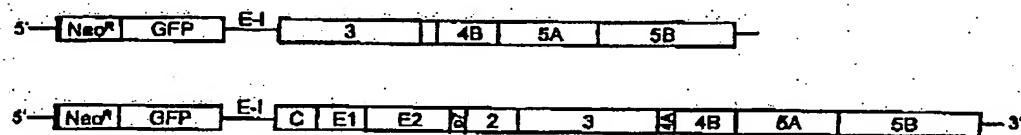
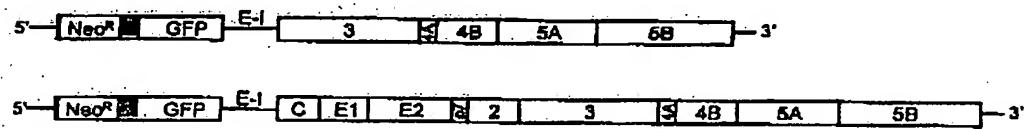
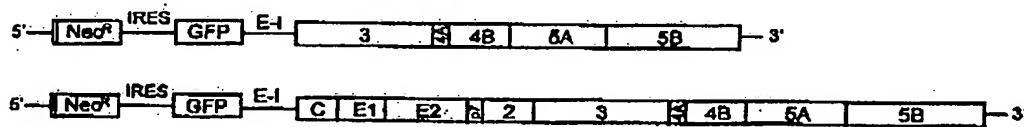


Fig. 5

A**B****C****D****Fig. 6**

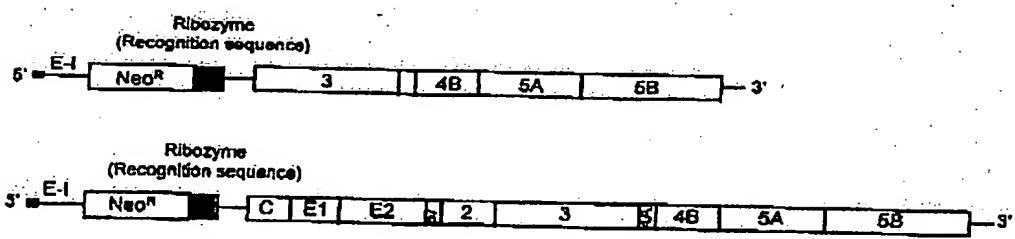


Fig. 7

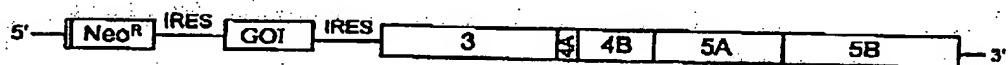


Fig. 8

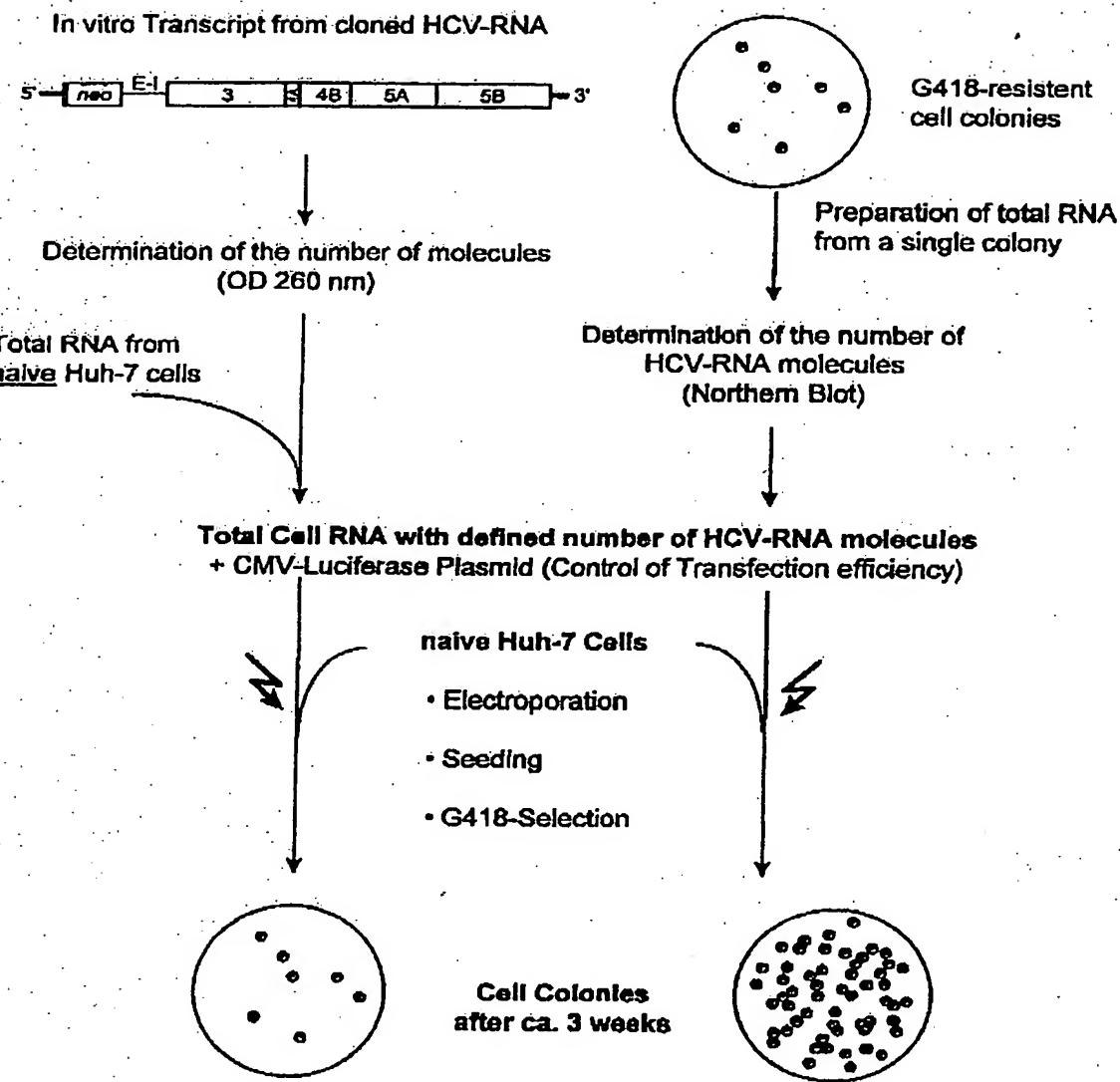


Fig. 9

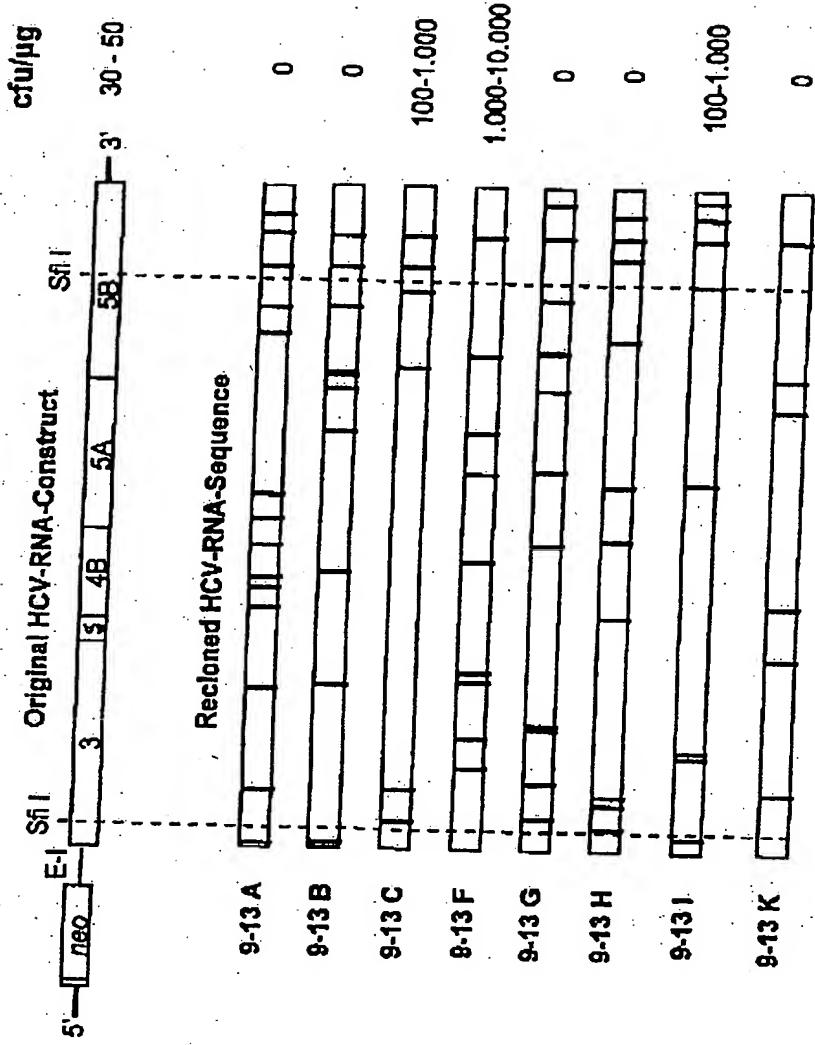
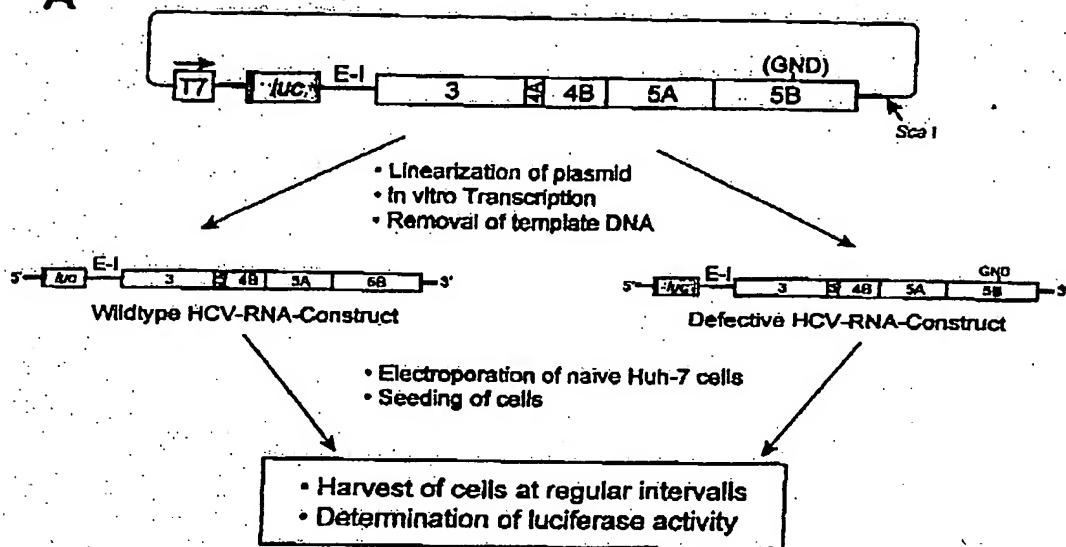
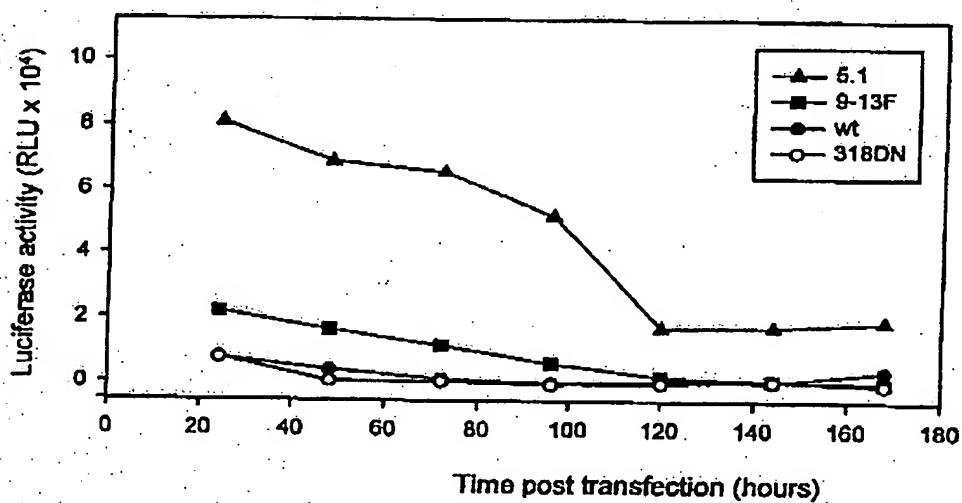


Fig. 10

Borden Ladner Gervais LLP / s.r.l.

A**B****Fig. 11**

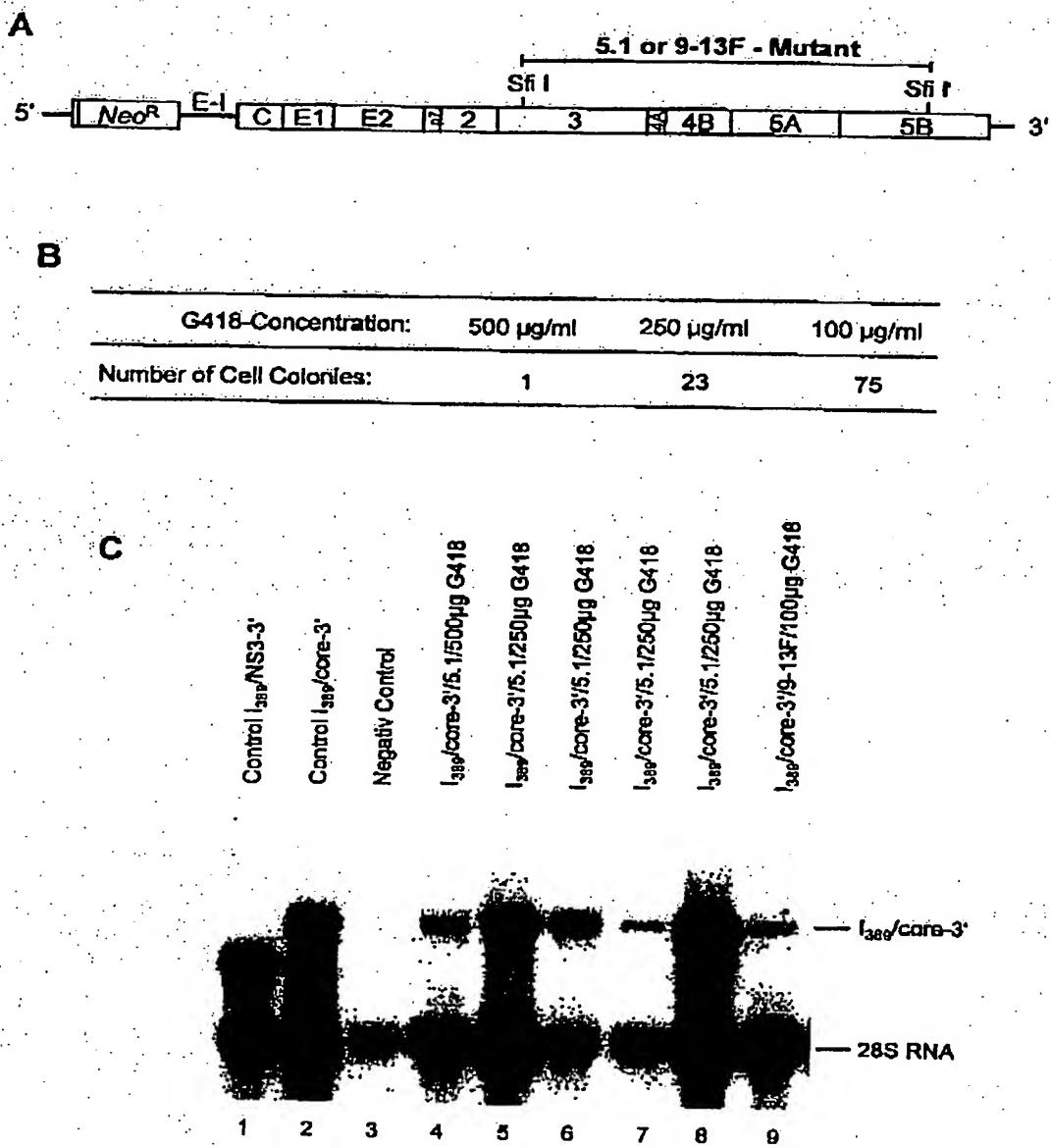


Fig. 12

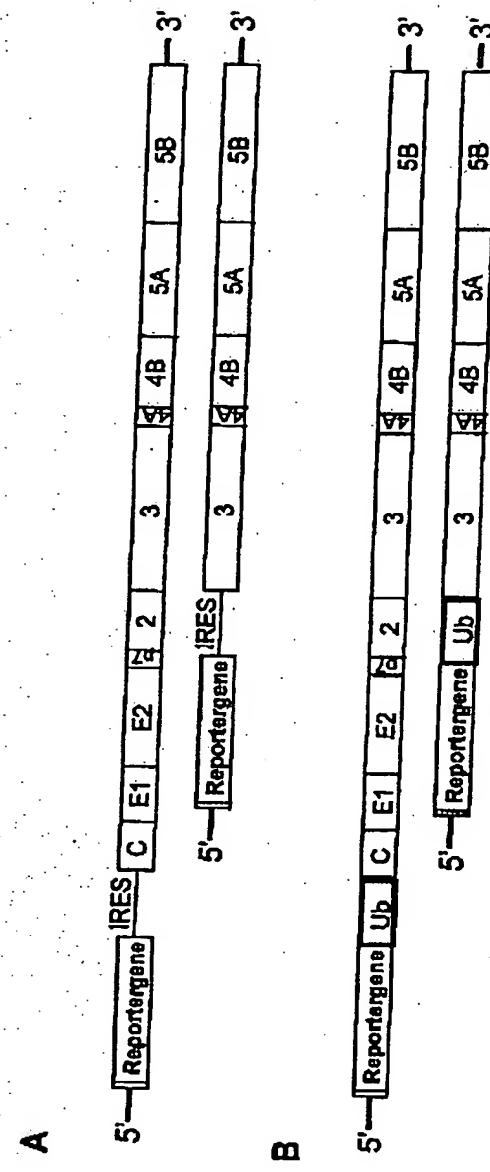


Fig. 12



Fig. 14

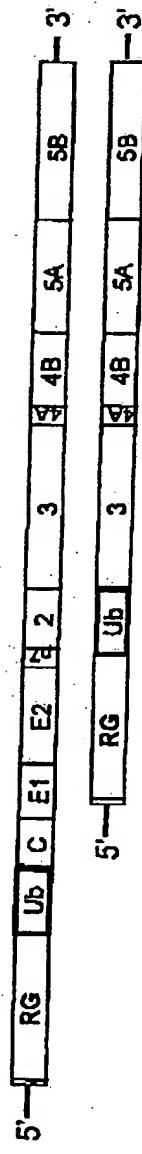


Fig. 15



CA 02303518 1000-03-31

Hepatitis C Virus Cell Culture SystemDESCRIPTION

5 The invention relates to a hepatitis C virus (HCV) cell culture system, which comprises mainly eukaryotic cells containing transfected HCV specific genetic material, which means they are transfected with HCV specific genetic material.

10

The hepatitis C virus (HCV) is one of the main causes worldwide of chronic and sporadic liver diseases. The history of most HCV infections does not involve any obvious clinical signs, but 80 - 90 % of the infected people become 15 chronic carriers of the virus and 50 % of these chronic carriers of the virus develop chronic hepatitis with different degrees of severity. Approx. 20 % of the chronically infected develop a cirrhosis of the liver over 10 to 20 years, based on what a primary hepatocellular 20 carcinoma can develop. Nowadays chronic hepatitis C is the main indication for liver transplantation. A specific therapy does not exist until now. The only therapy currently available is high-dose administration of Interferon alpha or a combination of Interferon alpha and 25 the purine nucleoside analogue Ribavirin. However, only approx. 60 % of all treated persons respond to this therapy and with these, a new viraemia occurs in more than half of all cases after the discontinuation of the treatment.

30 Due to the high prevalence, especially in industrialized countries, the serious effects of chronic infections and the non-existence of a specific therapy, the development of a HCV specific chemotherapy is an important goal of pharmaceutical research and development. The main problem 35 lies in the previous lack of a suitable cell culture system, which enables the study of virus replication and pathogenesis in eukaryotic cells.



Hepatitis C Virus Cell Culture System**DESCRIPTION**

5 The invention relates to a hepatitis C virus (HCV) cell culture system, which comprises mainly eukaryotic cells containing transfected HCV specific genetic material, which means they are transfected with HCV specific genetic material.

10

The hepatitis C virus (HCV) is one of the main causes worldwide of chronic and sporadic liver diseases. The history of most HCV infections does not involve any obvious clinical signs, but 80 - 90 % of the infected people become 15 chronic carriers of the virus and 50 % of these chronic carriers of the virus develop chronic hepatitis with different degrees of severity. Approx. 20 % of the chronically infected develop a cirrhosis of the liver over 10 to 20 years, based on what a primary hepatocellular 20 carcinoma can develop. Nowadays chronic hepatitis C is the main indication for liver transplantation. A specific therapy does not exist until now. The only therapy currently available is high-dose administration of Interferon alpha or a combination of Interferon alpha and 25 the purine nucleoside analogue Ribavirin. However, only approx. 60 % of all treated persons respond to this therapy and with these, a new viraemia occurs in more than half of all cases after the discontinuation of the treatment.

30 Due to the high prevalence, especially in industrialized countries, the serious effects of chronic infections and the non-existence of a specific therapy, the development of a HCV specific chemotherapy is an important goal of pharmaceutical research and development. The main problem lies in the previous lack of a suitable cell culture 35 system, which enables the study of virus replication and pathogenesis in eukaryotic cells.

Due to the small amount of virus in blood or tissue, the lack of suitable cell culture systems or animal models (the chimpanzee is still the only possible experimental animal) as well as the lack of efficient systems for producing virus-like particles, it was not possible up to now, to analyze the molecular composition of the HCV particle in-depth and to solve it. The information currently available can be summarized as follows: HCV is an enveloped plus-strand RNA virus with a particle diameter of 50 - 60 nm and a medium density of 1.03 - 1.1 g/ml. It was molecularly cloned and characterized for the first time in 1989 (Choo et al., 1989, *Science*, 244, 359 - 362). The HCV-RNA has a length of approx. 9.6 kb (= 9600 nucleotides), a positive polarity and comprises one open reading frame (ORF), which encodes a linear polyprotein of approx. 3010 amino acids (see Rice 1996, in *Virology*, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp. 931 - 960; Clarke 1997, *J. Gen. Virol.* 78, 2397; and Bärtenschlager 1997, *Intervirology* 40, 378 and see Fig. 1 A). During the replication of the virus the polyprotein is cleaved into the mature and functionally active proteins by cellular and viral proteases.

Within the polyprotein the proteins are arranged as follows (from the amino- to the carboxy terminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. The core protein is the main component of the nucleocapsid. The glycoproteins E1 and E2 are transmembrane proteins and the main components of the viral envelope. They probably play an important role during the attachment of the virus to the host cell. These three proteins core, E1, and E2 constitute the viral particle and are therefore called structural proteins. The function of the protein p7 is still not clear. The protein NS2 is probably the catalytic domain of the NS2-3 protease, which is responsible for the processing between the proteins NS2 and NS3. The protein NS3 has two functions, one is a

protease activity in the amino terminal domain, which is essential for the polyprotein processing, and the other a NTPase/helicase function in the carboxy terminal domain, which is probably important during the replication of the 5 viral RNA. The protein NS4A is a co-factor of the NS3 protease. The function of the protein NS4B is unknown.

The open reading frame is flanked on its 5' end by a non-translated region (NTR) approx. 340 nucleotides in length, 10 which functions as the internal ribosome entry site (IRES), and on its 3' end by a NTR approx. 230 nucleotides in length, which is most likely important for the genome replication. A 3' NTR such as this is the object of patent application PCT/US 96/14033. The structural proteins in the 15 amino terminal quarter of the polyprotein are cleaved by host cell signal peptidase. The non-structural proteins (NS) 2 to (NS) 5B are processed by two viral enzymes, namely the NS2-3 and the NS3/4A protease. The NS3/4A protease is required for all cleavages beyond the carboxy 20 terminus of NS3. The function of NS4B is unknown. NS5A, a highly phosphorylated protein, seems to be responsible for the Interferon resistance of various HCV genotypes (see Enomoto et al. 1995, *J. Clin. Invest.* 96, 224; Enomoto et al. 1996, *N. Engl. J. Med.* 334, 77; Gale Jr. et al. 1997, 25 *Virology* 230, 217; Kaneko et al. 1994, *Biochem. Biophys. Res. Commun.* 205, 320; Reed et al., 1997, *J. Virol.* 71, 7187), and NS5B has been identified as the RNA-dependent RNA polymerase.

30 First diagnostic systems have been developed from these findings, which are either based on the detection of HCV specific antibodies in patient serum or the detection of HCV specific RNA using the reverse transcription polymerase chain reaction (RT-PCR), and which are (must be) routinely used with all blood and blood products and/or according to 35 the regulations.

Since the first description of the genome in 1989 several partial and complete sequences of the HCV have been cloned and characterized using the PCR method. A comparison of these sequences shows a high variability of the viral genome in particular in the area of the NS5B gene, which eventually resulted in the classification of 6 genotypes, which are again subdivided into the subtypes a, b, and c.

The genomic variance is not evenly distributed over the genome. The 5'NTR and parts of the 3'NTR are highly conserved, while certain encoded sequences vary a lot, in particular the envelope proteins E1 and E2.

The cloned and characterized partial and complete sequences of the HCV genome have also been analyzed with regard to appropriate targets for a prospective antiviral therapy. In the course of this, three viral enzymes have been discovered, which may provide a possible target. These include (1) the NS3/4A protease complex, (2) the NS3 Helicase and (3) the NS5B RNA-dependent RNA polymerase. The NS3/4A protease complex and the NS3 Helicase have already been crystallized and their three-dimensional structure determined (Kim et al., 1996, Cell, 87, 343; Yen et al., 1998, Protein Science, 7, 837; Love et al., 1996, Cell, 87, 311; Kim et al., 1998, Structure, 6, 89; Yao et al., 1997, Nature Structural Biology, 4, 463, Cho et al., 1998, J. Biol. Chem., 273, 15045). It has not been successful until now with the NS5B RNA-dependent RNA polymerase.

Even though important targets for the development of a therapy for chronic HCV infection have been defined with these enzymes and even though a worldwide intensive search for suitable inhibitors is ongoing with the aid of rational drug design as well as high throughput screening, the development of a therapy has one major deficiency, namely the lack of cell culture systems or simple animal models, which allow direct, reliable identification of HCV-RNA or

HCV antigens with simple methods which are common in the laboratory. The lack of these cell culture systems is also the main reason that to date the comprehension of HCV replication is still incomplete and mainly hypothetical.

5

Although according to the experts, a close evolutionary relationship exists between HCV and the flavivirus and pestiviruses, and self-replicating RNAs have been described for these, which can be used for the replication in different cell lines with a relatively high yield, (see Khromykh et al., 1997, *J. Virol.* 71, 1497; Behrens et al., 1998, *J. Virol.* 72, 2364; Moser et al., 1998, *J. Virol.* 72, 5318), similar experiments with HCV have not been successful to date.

15

Although it is known from different publications that cell lines or primary cell cultures can be infected with high titre patient serum containing HCV, (Lanford et al. 1994, *Virology* 202, 606; Shimizu et al. 1993, *Proceedings of the National Academy of Sciences, USA*, 90, 6037 - 6041; Mizutani et al. 1996, *Journal of Virology*, 70, 7219 - 7223; M. Ikeda et al. 1998, *Virus Res.* 56, 157; Fournier et al. 1998, *J. Gen. Virol.* 79, 2376 and bibliographical references quoted in here; Ito et al. 1996, *Journal of General Virology*, 77, 1043 - 1054), these virus-infected cell lines or cell cultures do not allow the direct detection of HCV-RNA or HCV antigens. The viral RNA in these cells can not be detected in a Northern Blot (a standard method for the quantitative detection of RNA) or the viral protein in a Western Blot or with immunoprecipitation. It has only been possible to detect HCV replication with very costly and indirect methods. These disadvantageous facts show that obviously the replication in these known virus-infected cell lines or cell cultures is completely insufficient.

Furthermore it is known from the publications of Yoo et al. (1995, *Journal of Virology*, 69, 32 - 38) and of Dash et al., (1997, *American Journal of Pathology*, 151, 363 - 373) that hepatoma cell lines can be transfected with synthetic HCV-RNA, which are obtained through in vitro transcription of the cloned HCV genome. In both publications the authors started from the basic idea that the viral HCV genome is a plus-strand RNA functioning directly as mRNA after being transfected into the cell, permitting the synthesis of viral proteins in the course of the translation process, and so new HCV particles are (could be) formed. This viral replication, which means these newly formed HCV viruses and their RNA, have been detected through RT-PCR. However the published results of the RT-PCR carried out indicate, that the HCV replication in the described HCV transfected hepatoma cells is not particularly efficient and is not sufficient to measure the quality, let alone the quantity of the fluctuations in the replication rate after an targeted action with prospective antiviral treatments.

Furthermore it is prior art (Yanagi et al., Proc. Natl. Acad. Sci. USA, 96, 2291-95, 1999), that the highly conserved 3' NTR is essential for the virus replication. This knowledge strictly contradicts the statements of Yoo et al. and Dash et al., who used for their experiments only HCV genomes with shorter 3' NTRs since they did not know the authentic 3' end of the HCV genome.

The object of the present invention is to provide a HCV cell culture system, where the viral RNA self-replicates in the transfected cells with such a high efficiency that the quality and quantity of the fluctuations in the replication rate can be measured with common methodologies usually found in the laboratory after a targeted action with virus and prospective HCV specific antivirals in particular.

35

The solution to this problem is to provide a cell culture system of the prior mentioned type, where the eukaryotic

cells are human cells, in particular hepatoma cells, which are preferably derived from a normal hepatoma cell line, but can also be obtained from an appropriate primary cell culture, and where the transfected HCV specific genetic material is a HCV-RNA construct, which essentially comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR preferably in the order mentioned as well as a minimum of one marker gene for selection (selection gene).

10

Here and in the following "NTR" stands for "non-translated region" and is a known and familiar term or abbreviation to the relevant expert.

15

Here and in the following the term "HCV-RNA construct" comprises constructs, which include the complete HCV genome, as well as those, which only include a part of it, which means a HCV subgenome.

20

A preferred variation of the cell culture system according to the invention, which had proven to be worthwhile in practice, is lodged at the DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (German collection of Microorganisms and Cell Cultures) in Braunschweig, Germany under the number DSM ACC2394 (laboratory name HuBl 9-13).

With the cell culture system according to the invention an in vitro system is provided for the first time, where HCV-RNA is self-replicated and expressed intracellularly and in a sufficient amount, so that the quantity of the amounts of HCV-RNA as well as the HCV specific proteins can be determined with conventional and reliably precise biochemical measuring methods. This means an almost authentic cell-based HCV replication system is available for the first time, which is urgently needed for the development and testing of antiviral drugs. This test system provides the possibility of identifying potential

targets for an effective HCV specific therapy and developing and evaluating HCV specific chemotherapeutics.

The invention is based on the surprising finding that efficient replication of the HCV-RNA only occurs in cells if they have been transfected with an HCV-RNA construct, which comprises at least the 5' and the 3' non-translated regions (NTR) and the non-structural proteins (NS) 3 to 5B and additionally a marker gene for selection (selection gene). The structural genes are obviously without great importance for replication, whereas efficient replication of the HCV-RNA apparently only occurs if the transfected cells are subject to permanent selection pressure, which is imparted by the marker gene for selection (selection gene) linked to the HCV-RNA. Consequently the marker gene (selection gene) seems on one hand to provoke the selection of those cells, where the HCV-RNA replicates productively, and it seems on the other hand to considerably increase the efficiency of the RNA replication.

20

An object of the invention is also a cell-free HCV-RNA construct, characterized in that it comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR, preferably in the order mentioned, as well as a 25 marker gene for selection (selection gene).

30

In the present context the terms 5' NTR and NS3 and NS4A and NS4B and NS5A and NS5B and 3' NTR comprise each nucleotide sequence, which is described in the state of the art as the nucleotide sequence for each functional segment of the HCV genome.

35

By providing a HCV-RNA construct such as this, a detailed analysis of the HCV replication, pathogenesis and evolution in cell culture is possible for the first time. The HCV specific viral RNA can specifically be created as a complete genome or subgenome in any amount, and it is

possible, to manipulate the RNA construct and consequently to examine and identify the HCV functions on a genetic level.

5 Because all HCV enzymes identified as a main target for a therapy at the moment, namely the NS3/4A protease, the NS3 helicase and the NS5B polymerase, are included in the HCV-RNA construct according to the invention, it can be used for all relevant analyses.

10

An embodiment of the HCV-RNA construct, which has proven to be worthwhile in practical use, stands out by the fact that it comprises the nucleotide sequence according to the sequence protocol SEQ ID NO:1.

15

Further embodiments with similar good properties for practical use are characterized in that they comprise a nucleotide sequence either according to sequence protocol SEQ ID NO:2 or according to sequence protocol SEQ ID NO:3 or according to sequence protocol SEQ ID NO:4 or according to sequence protocol SEQ ID NO:5 or according to sequence protocol SEQ ID NO:6 or according to sequence protocol SEQ ID NO:7 or according to sequence protocol SEQ ID NO:8 or according to sequence protocol SEQ ID NO:9 or according to sequence protocol SEQ ID NO:10 or according to sequence protocol SEQ ID NO:11.

30 It is possible to provide the HCV subgenomic construct with a 3' NTR, which has a nucleotide sequence so far unknown in the state of art, a nucleotide sequence, which has been selected from the group of nucleotide sequences (a) to (i) listed in the following:

(a) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTAGCTTT
35 TTTTTTTTCTTTTTTGAGAGAGAGTCTCACTCTGTTGCCAGACTGGAG
T

(b) ACAGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTCT
TTTTTTTTTC TTTTTTTGA GAGAGAGAGT CTCACTCTGT TGCCCAGACT
GGAGC

5 (c) ACAGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTAATCTTT
TTTTTTTTCT TTTTTTTGA GAGAGAGAGT CTCACTCTGT TGCCCAGACT
GCAGC

10 (d) ACAGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTC
TTTTTTTTT TCTTTTTTT TGAGAGAGAG AGTCTCACTC TGTTGCCAG
ACTGGAGT

15 (e) ACAGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTCT
TTTTTTTTT TCTTTTTTT TGAGAGAGAG AGTCTCACTC TGTTGCCAG
ACTGGAGT

(f) ACAGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTCT
TTTTTTTTT TCTTTTTTT TGAGAGAGAGA GAGTCTCACT CTGTTGCCAG
GACTGGAGT

20 (g) ACAGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTCT
TTTTTTTTT CTTTTTTTT GAGAGAGAGA GTCTCACTCT GTTGCCAGA
CTGGAGT

25 (h) ACAGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTTAAT
CTTTTTTTT TTTTCCTTT TTTGAGAGA GAGAGTCTCA CTCTGTTGCC
CAGACTGGAG T

30 (i) ACAGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTTAATC
TTTTTTTTT TTTCTTTT TTTTGAGAG AGAGAGTCTC ACTCTGTTGC
CCAGACTGGA GT

The marker gene for selection (selection gene) included in
the HCV-RNA constructs according to the invention is
35 preferably a resistance gene, in particular an antibiotic
resistance gene.

5 This has the advantage that the cells transfected with this construct can easily be selected from the non-transfected cells by adding for example the appropriate antibiotic to the cell culture medium in the case of an antibiotic resistance gene.

In the present context 'antibiotic' means any substance, which impedes the non-transfected host cells or the cells, where the HCV-RNA is not replicating efficiently,
10 continuing to live or grow, especially the cell poison Puromycin, Hygromycin, Zeocin, Bleomycin or Blasticidin.

15 A preferred marker gene for selection (selection gene) and resistance gene, which has proven to be worthwhile in practice, is the neomycin phosphotransferase gene.

An alternative for the antibiotic resistance genes is for example the thymidine kinase gene, which can be used to carry out a HAT selection.

20

The marker gene for selection (selection gene), the preferred resistance gene and the most preferred antibiotic resistance gene is preferably positioned in the HCV-RNA construct after the HCV 5' NTR, which means downstream from the 5' NTR and upstream from the HCV reading frame.
25 However, an insertion in the area of the 3' NTR or another site of the HCV genome or subgenome, for example within the polyprotein, is also conceivable.

30 In another embodiment of the HCV-RNA construct according to the invention the marker gene for selection (selection gene), in particular an antibiotic resistance gene, is linked to the HCV-RNA or HCV genomic or subgenomic sequence via a ribozyme or a recognition site for a ribozyme.

35

This has the advantage, that after the selection of the cells, in which the HCV-RNA is replicating productively,

the resistance gene in the obtained cell clones can be separated from the HCV subgenomic sequence through a ribozyme-dependent cleavage, namely by activating the inserted ribozyme or in the case of a construct with a recognition site for a ribozyme, by transfecting the ribozyme into the cells (for example through the transfection of a ribozyme construct or infection with a viral expression vector, into which the appropriate ribozyme has been inserted). By this means an authentic HCV genomic construct can be obtained without a resistance gene, which can then form authentic infectious virus particles.

Another preferred embodiment of the HCV-RNA construct is according to the invention is characterized in that the construct has at least one integrated reporter gene.

In the following a reporter gene means any gene, whose presence can be easily detected with, in general, simple biochemical or also histochemical methods after being transferred into a target organism, which means a gene, that encodes for a protein, which can be easily and reliably detected and quantified in small amounts with the common measuring methods in the laboratory.

25

This variation of the HCV-RNA construct has the advantage that the extent of the replication of this construct can be easily and quickly measured with the methods common in the laboratory using the reporter gene product.

30

The reporter gene is preferably a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) or the SEAP gene (secreted alkaline phosphatase gene). This reporter gene and its products, namely the relevant reporter proteins, can be

detected for example using fluorescence, chemiluminescence, colorimetric measurements or by means of immunological methods (for example ELISA).

5 A surrogate marker gene can also be considered as a reporter gene. In this context it includes those genes, which encode for cellular proteins, nucleic acids or generally for those functions, which are subject to variation depending on the replication of the virus, and
10 which consequently are either suppressed or activated in the cells, in which the HCV or the HCV-RNA construct multiplies. This means, the suppression or activation of this function is a surrogate marker for the replication of the virus or the replication of the HCV-RNA construct.

15

The positions of the reporter genes and the marker gene for selection (selection gene) can be selected in such a way, that a fusion protein made from both genetic products will be expressed. This has the advantage that these two genes
20 can be arranged in such a way in the HCV-RNA construct that their two expressed proteins are fused via a recognition sequence for a protease (for example ubiquitin) or via a self-cleaving peptide (for example the 2A protein of the Picornaviruses) at first and will be separated
25 proteolytically later.

These two positions might as well lie apart from each other in such way, that both genetic products are separately expressed (for example in the order: marker or resistance
30 gene - internal ribosome binding site - reporter gene).

In the case of the reporter gene one embodiment has proven to be particularly worthwhile, where the reporter gene is cloned into the open reading frame of the HCV genome or
35 subgenome in such a way that it will only be transferred to an active form after proteolytic processing.

The cell culture system according to the invention can be used for various purposes in each of its embodiments. These comprise:

- 5 • The detection of antiviral substances. This can include for example: organic compounds, which interfere directly or indirectly with viral growth (for example inhibitors of the viral proteases, the NS3 helicase, the NS5B RNA-dependent RNA polymerase), antisense oligonucleotides,
10 which will hybridize to any target sequence in the HCV-RNA construct (for example the 5' NTR) and will have an direct or indirect influence on the virus growth for example due to a reduction of translation of the HCV polyprotein or ribozymes, which cleave any HCV-RNA
15 sequence and consequently impair virus replication.
- The evaluation of any type of antiviral substances in the cell culture. These substances can be detected on the isolated purified enzyme for example with 'rational drug design' or 'high-throughput screening'. Evaluation means mainly the determination of the inhibitory features of the respective substance as well as its mode of action.
20
- 25 • The identification of new targets of viral or cellular origin for a HCV specific antiviral therapy. If for example a cellular protein is essential for viral replication, the viral replication can also be influenced by inhibiting this cellular protein. The system according to the invention also enables the detection of these auxiliary factors.
30
- 35 • The determination of drug resistance. It can be assumed that resistance to therapy occurs due to the high mutation rate of the HCV genome. This resistance, which is very important for the clinical approval of a substance, can be detected with the cell culture system.

according to the invention. Cell lines, in which the HCV-RNA construct or the HCV genome or subgenome replicates, are incubated with increasing concentrations of the relevant substance and the replication of the viral RNA is either determined by means of an introduced reporter gene or through the qualitative or quantitative detection of the viral nucleic acids or proteins. Resistance is given if no or a reduced inhibition of the replication can be observed with the normal concentration of the active substance. The nucleotide and amino acid replacements responsible for the therapy resistance can be determined by recloning the HCV-RNA (for example by the means of RT-PCR) and sequence analysis. By cloning the relevant replacement/s into the original construct its causality for the resistance to therapy can be proven.

- The production of authentic virus proteins (antigens) for the development and/or evaluation of diagnostics.
- The cell culture system according to the invention also allows the expression of HCV antigens in cell cultures. In principle these antigens can be used as the basis for diagnostic detection methods.
- The production of HCV viruses and virus-like particles, in particular for the development or production of therapeutics and vaccines as well as for diagnostic purposes. Especially cell culture adapted complete HCV genomes, which could be produced by using the cell culture system according to the invention, are able to replicate in cell culture with high efficiency. These genomes have the complete functions of HCV and in consequence they are able to produce infectious viruses.
- The HCV-RNA construct according to the invention by itself can also be used for various purposes in all its embodiments. This includes first of all:

- The construction of attenuated hepatitis C viruses or HCV-like particles and their production in cell cultures:

5

10

20

25

30

35

Attenuated HCV or HCV-like particles can be created by accidental or purposefully introduced mutations, such as point mutations, deletions or insertions, which means viruses or virus-like particles with complete ability to replicate, but reduced or missing pathogenicity. These attenuated HCV or HCV-like particles can be used in particular as vaccine.

• The construction of HCV-RNA constructs with integrated foreign genes, used for example as liver cell specific vector in gene therapy. Due to the distinctive liver cell tropism of the HCV and the possibility of replacing parts of the genome by heterologous sequences, HCV-RNA constructs can be produced, where for example the structural proteins can be replaced by a therapeutically effective gene. The HCV-RNA construct obtained in this way is introduced into cells preferably by means of transfection, which express the missing HCV functions, for example the structural proteins, in a constitutive or inducible way. Virus particles, carrying the HCV-RNA construct, can be created by means of this method known to the expert under the term 'transcomplementation'. The particles obtained can preferably be used for the infection of liver cells. Within these the therapeutically effective foreign gene will be expressed and will consequently develop its therapeutic effect.

• The detection of permissive cells, which means cells, in which a productive virus growth occurs. For this purpose either one of the HCV-RNA genomic constructs previously mentioned, which is able to form complete infectious viruses, or one of the HCV subgenomes previously

mentioned, which according to the previously mentioned example will be transfected in a cell line first, which expresses the missing functions in a constitutive or inducible way, is used. In each case virus particles are created, which carry a resistance and/or reporter gene apart from the HCV sequence. In order to detect cells, where the HCV is able to replicate, these cells are infected with viruses generated in this way and subject to an antibiotic selection or they are examined depending on the HCV-RNA construct by means of determining the presence of the expression of the reporter gene. Because an antibiotic resistance or reporter gene expression can only be established, when HCV-RNA construct replicates, the cells detected in this way must be permissive. Almost any cell line or primary cell culture can be tested in regard to the permissivity and detected in this way.

- 20 The cell culture system according to the invention also permits targeted discovery of HCV-RNA constructs for which there is an increase in the efficiency of replication due to mutations. This occurs either by chance, in the context of HCV-RNA replication, or by targeted introduction into
25 the construct. These mutations, leading to a change in the replication of the HCV-RNA construct, are known to experts as adaptive mutations. The invention therefore also includes a method for obtaining cell culture adapted mutants of a HCV-RNA construct according to the invention
30 following the above description, in which the mutants have increased replication efficiency compared to the original HCV-RNA construct. It further includes a method for the production of mutants of a HCV-RNA full-length genome or of a HCV-RNA subgenome or of any HCV-RNA construct with
35 increased replication efficiency compared to the original HCV-RNA full-length genome or subgenome or HCV-RNA construct, as well as cell culture adapted mutants of HCV-

RNA constructs, HCV-RNA full-length genomes and HCV subgenomes with increased replication efficiency compared to the original constructs, subgenomes or full-length genomes.

5

The method according to the invention for the production of cell culture adapted mutants of a HCV-RNA construct according to the invention, in which the mutants have increased replication efficiency compared to the HCV-RNA construct, is characterised in that a cell culture system according to claim 1, in which the transfected HCV specific genetic material is a HCV-RNA construct with a selection gene according to claims 4 to 19, is cultivated on/in the selection medium corresponding to the selection gene, that the cultivated cell clones are collected and that the HCV-RNA construct is isolated from these cell clones.

In an advantageous extension of this production method, the isolated HCV-RNA constructs are passaged at least one more time, that is they are transfected in cells of a cell culture system according to claim 1, the thus obtained cell culture system according to claim 1, in which the transfected HCV specific genetic material is the isolated HCV-RNA construct with a selection gene, is cultivated on/in the selection medium corresponding to the selection gene, the cultivated cell clones are collected and the HCV-RNA constructs are thus isolated.

30

Using this process variation, the quantity of adaptive mutations and hence the degree of replication efficiency in the relevant HCV-RNA constructs can be increased even further.

35

The method according to the invention for the production of mutants of a HCV-RNA full-length genome or of a HCV-RNA subgenome or of any HCV-RNA construct with increased replication efficiency compared to the original HCV-RNA

full-length genome or subgenome or HCV-RNA construct, has the following features. Using one of the two production methods presented above, a cell culture adapted mutant of a HCV-RNA construct is produced, isolated from the cells, 5 cloned using state of the art known methods and sequenced. By comparing with the nucleotide and amino acid sequence of the original HCV-RNA construct, the type, number and position of the mutations is determined. These mutations are then introduced into an (isolated) HCV subgenome or 10 full-length genome or any HCV-RNA construct, either by site-directed mutagenesis, or by exchange of DNA fragments containing the relevant mutations.

A test can be carried out to determine or verify which 15 mutations actually are responsible for an alteration of replication efficiency, particularly an increase in replication. In this test the corresponding nucleotide and/or amino acid changes are introduced into the original HCV-RNA construct and the modified construct is then 20 transfected in cell culture. If the introduced mutation actually leads to an increase in replication, then for a HCV-RNA construct with a selectable marker gene, the number of resistant cell clones in the artificially mutated construct should be noticeably higher compared to the 25 untreated construct.

In the case of a construct with a reporter gene, the activity or quantity of the reporter should be noticeably higher for the artificially mutated construct compared to 30 the untreated one.

The cell culture adapted HCV-RNA constructs with high replication efficiency according to the invention are characterized in that, through nucleotide or amino acid 35 exchanges, they are derivable from a HCV-RNA construct according to one of the claims 4 to 19, and that they are

obtainable using one of the two production processes presented above.

These cell culture adapted HCV-RNA constructs can be used to produce any HCV-RNA constructs or HCV full-length or subgenomes with increased replication efficiency. Both constructs with a selectable resistance gene and constructs without one or with a non-selectable reporter gene (e.g. luciferase) can be produced in this way, since replication of cell culture adapted HCV-RNA constructs can also be demonstrated in non-selected cells due to their high replication efficiency.

The cell culture adapted mutants of a HCV-RNA construct or HCV-RNA full-length genome or HCV subgenome with high replication efficiency compared to the original HCV-RNA construct or the original HCV full-length genome are characterized in that they are obtainable by a method in which the type and number of mutations in a cell culture adapted HCV-RNA construct are determined through sequence analysis and sequence comparison and these mutations are introduced into a HCV-RNA construct, particularly a HCV-RNA construct according to one of the claims 4 to 19, or into an (isolated) HCV-RNA full-length genome, either by site-directed mutagenesis, or by exchange of DNA fragments containing the relevant mutations.

A group of preferred HCV-RNA constructs, HCV full-length genomes and HCV subgenomes with high and very high replication efficiency, which are consequently highly suitable for practical use is characterised in that it contains one, several or all of the amino acid or nucleic acid exchanges listed in table 3 and/or one or several of the following amino acid exchanges: 1283 arg -> gly , 1383 glu -> ala , 1577 lys -> arg , 1609 lys -> glu , 1936 pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile

-> val. (The numbers refer to the amino acid positions of the polyprotein of the HCV isolate con1, see Table 1).

5 Special features of the nucleotid sequences according to the sequence listings:

SEQ ID-NO: 1

Name: I389/Core-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 10 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 15 4.1813-10842: HCV Polyprotein from Core up to nonstructural protein 5B
 5. 1813-2385: HCV Core Protein; structural protein
 6. 2386-2961: envelope protein 1 (E1); structural protein
 - 20 7. 2962-4050: envelope protein 2 (E2); structural protein
 8. 4051-4239: Protein p7
 9. 4240-4890: nonstructural protein 2 (NS2); HCV NS2-3 Protease
 - 25 10.4891-6783: nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
 - 11.6784-6945: nonstructural protein 4A (NS4A); NS3 Protease cofactor
 - 12.6946-7728: nonstructural protein 4B (NS4B)
 - 30 13.7729-9069: nonstructural protein 5A (NS5A)
 - 14.9070-10842: nonstructural protein 5B (NS5B); RNA-dependent RNA-polymerase
 - 15.10846-11076: HCV 3' non-translated region

SEQ ID-NO: 2

Name: I337/NS2-3'/wt

Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

5 2.342-1181: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker

3.1190-1800: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame

10 4.1801-8403: HCV Polyprotein from nonstructural protein 2 up to nonstructural protein 5B

5. 1801-2451: nonstructural protein 2 (NS2); HCV NS2-3 Protease

6. 2452-4344: nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase

7. 4345-4506: nonstructural protein 4A (NS4A); NS3 Protease cofactor

8. 4507-5289: nonstructural protein 4B (NS4B)

9. 5290-6630: nonstructural protein 5A (NS5A)

20 10.6631-8403: nonstructural protein 5B (NS5B); RNA-dependent RNA-polymerase

11.8407-8637: HCV 3' non-translated region

SEQ ID-NO: 3

25 Name: I389/NS3-3'/wt

Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker

30 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame

4.1813-7767: HCV Polyprotein from nonstructural protein 3 up to nonstructural protein 5B

- 23 -

5. 1813-3708: nonstructural protein 3 (NS3); HCV NS3
Protease/Helicase
6. 3709-3870: Nonstructural protein 4A (NS4A); NS3
Protease Cofactor
- 5 7. 3871-4653: Nonstructural protein 4B (NS4B)
8. 4654-5994: Nonstructural protein 5A (NS5A)
9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
10. 7771-8001: HCV 3' non-translated Region

10

SEQ ID-NO: 4

Name: I337/NS3-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
15. 2.342-1181: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1190-1800: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
20. 4.1801-7758: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B
5. 1801-3696: Nonstructural protein 3 (NS3); HCV NS3
Protease/Helicase
6. 3697-3858: Nonstructural protein 4A (NS4A); NS3
25 Protease Cofactor
7. 3859-4641: Nonstructural protein 4B (NS4B)
8. 4642-5982: Nonstructural protein 5A (NS5A)
9. 5983-7755: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
30. 10. 7759-7989: HCV 3' non-translated Region

SEQ ID-NO: 5

Name: I389/NS2-3'/wt

Composition (Nucleotide positions):

- 24 -

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1813-8418: HCV Polyprotein from Nonstructural protein 2 up to Nonstructural protein 5B
5. 1813-2463: Nonstructural protein 2 (NS2); HCV NS2-3 Protease
6. 2464-4356: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
7. 4357-4518: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
8. 4519-5301: Nonstructural protein 4B (NS4B)
9. 5302-6642: Nonstructural protein 5A (NS5A)
10. 6643-8415: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
11. 8419-8649: HCV 3' non-translated Region

20

SEQ ID-NO: 6

Name: I389/NS3-3'/9-13F

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell culture-adapted mutant 9-13F
 5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3

Protease Cofactor

7. 3871-4653: Nonstructural protein 4B (NS4B)
 8. 4654-5994: Nonstructural protein 5A (NS5A)
 9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
- 5 7771-8001: HCV 3' non-translated Region

SEQ ID-NO: 7

Name: I389/Core-3'/9-13F

10 Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
 - 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
 - 3.1202-1812: internal ribosome entry site from
- 15 encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1813-10842: HCV Polyprotein from Core up to Nonstructural protein 5B of the cell culture-adapted mutant 9-13F
 5. 1813-2385: HCV Core Protein; structural protein
 6. 2386-2961: envelope protein 1 (E1); structural protein
 7. 2962-4050: envelope protein 2 (E2); structural protein
 8. 4051-4239: Protein p7

25 9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3 Protease

10.4891-6783: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase

11.6784-6945: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor

12.6946-7728: Nonstructural protein 4B (NS4B)

13.7729-9069: Nonstructural protein 5A (NS5A)

14.9070-10842: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase

35 15.10846-11076: HCV 3' non-translated Region

SEQ ID-NO: 8

Name: I389/NS3-3'/5.1

Composition (Nucleotide positions):

1. 1-341: HCV 5' non-translated region
5. 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
10. 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell culture-adapted mutant 5.1
5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
15. 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
7. 3871-4653: Nonstructural protein 4B (NS4B)
8. 4654-5994: Nonstructural protein 5A (NS5A)
9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
20. 7771-8001: HCV 3' non-translated Region

SEQ ID-NO: 9

Name: I389/Core-3'/5.1

25. Composition (Nucleotide positions):
- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
30. 4.1813-10842: HCV Polyprotein from Core up to Nonstructural protein 5B of the cell culture-adapted mutant 5.1
5. 1813-2385: HCV Core Protein; structural protein

- 27 -

6. 2386-2961: envelope protein 1 (E1); structural protein
7. 2962-4050: envelope protein 2 (E2); structural protein
8. 4051-4239: Protein p7
9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
11. 6784-6945: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
12. 6946-7728: Nonstructural protein 4B (NS4B)
13. 7729-9069: Nonstructural protein 5A (NS5A)
14. 9070-10842: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
15. 10846-11076: HCV 3' non-translated Region

SEQ ID-NO: 10

Name: I389/NS3-3'/19

20. Composition (Nucleotide positions):
 - 1.1-341: HCV 5' non-translated region
 - 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
 - 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
 - 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell culture-adapted mutant 19
30. 5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
6. 3709-3870: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
7. 3871-4653: Nonstructural protein 4B (NS4B)
35. 8. 4654-5994: Nonstructural protein 5A (NS5A)

9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase

7771-8001: HCV 3' non-translated Region

5 SEQ ID-NO: 11

Name: I389/Core-3'/19

Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

2.342-1193: HCV Core Protein-Neomycin Phosphotransferase

10 fusion protein; selectable Marker

3.1202-1812: internal ribosome entry site from
encephalomyokarditis virus; directs translation of the
downstream located HCV open reading frame

4.1813-10842: HCV Polyprotein from Core up to Nonstructural
protein 5B of the cell culture-adapted mutant 19

5. 1813-2385: HCV Core Protein; structural protein

6. 2386-2961: envelope protein 1 (E1); structural
protein

7. 2962-4050: envelope protein 2 (E2); structural
protein

8. 4051-4239: Protein p7

9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3
Protease

10.4891-6783: Nonstructural protein 3 (NS3); HCV NS3
Protease/Helicase

25 11.6784-6945: Nonstructural protein 4A (NS4A); NS3
Protease Cofactor

12.6946-7728: Nonstructural protein 4B (NS4B)

13.7729-9069: Nonstructural protein 5A (NS5A)

30 14.9070-10842: Nonstructural protein 5B (NS5B); RNA-
dependent RNA-Polymerase

15.10846-11076: HCV 3' non-translated Region

The invention is described in detail in the following by way of examples of embodiments and respective diagrams. These are as follows.

5 Fig. 1 A: The structure of a HCV-RNA construct according to the invention

10 On the top a diagram of the structure of the complete parental HCV genome is given with the positions of the genes for the cleavage products core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B within the polyprotein, and the 5' and 3' non-translated regions (5' NTR and 3' NTR) shown as a thick horizontal line -, and with the two positions selected for the creation of the 15 subgenomic constructs, namely the position of the 'GDD catalytic domain' of the NS5B RNA polymerase (GDD) and the position of the 3' end of the HCV-IRES (nucleotide positions 1 to 377 and 1 to 389) drawn above the diagram of the genome. The numbers below the diagram of the genome indicate the respective nucleotide positions.

20
25 Diagrams of the structure of two modified HCV-RNA constructs (subgenome) according to the invention are shown below, consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (Neo^R), the EMCV-IRES (E-I) and the HCV sequences of NS2 or NS3 up to the authentic 3' end. The position of the 10-amino acid deletion comprising the NS5B polymerase GDD motive is marked with a triangle (Δ).

30
35 Fig. 1 B: The result of a denaturing formaldehyde-agarose gel electrophoresis for the detection of replicating plus-strand RNA in transfected subpassaged Huh 7 cell clones.

The positions of HCV specific RNAs (arrows) and the 28S rRNA are specified to the right of lane 12, the size (number of nucleotides) of the RNA marker (M) is specified to the left of lane 1.

Fig. 1 C: The result of a PCR test with subsequent Southern Blot to demonstrate the absence of integrated replicon DNA in most of the selected cell clones.

The lanes 1 and 2 show the positive control, lane 13 the negative control. The figures to the left of lane 1 indicate the size of the nucleotide marker molecules.

Fig. 2 A: The result of a PCR test with subsequent Southern Blot for the detection of the sensitive exclusion of integrated replicon DNA (plasmid molecule I₃₇₇/NS3-3'/wt) in a cell clone containing a HCV-RNA construct (9-13).

The lanes 7 to 11 represent the result of a titration of DNA molecules of the construct I₃₇₇/NS3-3'/wt without addition of total DNA of the cell clone 9-13, and the lanes 2 - 6 represent the result obtained with the same plasmid molecules with the addition of 1 µg 9-13 DNA each prior to the PCR (for the purpose of excluding an inhibitor of the PCR in the DNA preparation). Lane 13 represents the negative control (PCR without DNA template). Lane 1 shows the result, which was achieved with one µg total DNA of the cell clone 9-13.

Fig. 2 B: The result of a Northern Blot test to quantify the amounts of HCV plus- and minus-strand RNA.

The arrows mark the positions of replicon RNA.

The "plus" and "minus" details indicate the positive (plus) and negative (minus) polarity of the RNA controls, which have been applied to the gel. "Minus-strand" and "Plus-strand" indicate the specificity of the radioactive RNA probes.

Fig. 2 C: The result of a formaldehyde-agarose gel electrophoresis after radioactive labeling of the intracellular replicated HCV-RNA to demonstrate the resistance of HCV-RNA replication to dactinomycin.

Fig. 3 A: The detection of HCV specific antigens in the selected cell clones by means of immunoprecipitation after metabolic radioactive labeling.

The lanes 7 - 9 represent authentic size marker (which have been obtained in Huh 7 cells after the transient expression of a HCV-RNA construct); identified HCV proteins are labeled on the left edge of lane 1, the molecular weights (in Kilodalton) are specified to the right of lane 9.

Fig. 3 B: Results of an immunofluorescence test to establish the subcellular localisation of HCV antigens.

Fig. 4 : Diagram of the structure of a selectable HCV-RNA construct according to the invention (complete genome) consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (NeoR), a heterologous IRES element, for example from the encephalomyocarditis virus (E-I), the complete HCV open reading frame and the authentic 3' NTR.

5 **Fig. 5:** Diagram of the structure of HCV-RNA constructs
with inserted antibiotic resistance gene (A)
within the nucleotide sequence encoding the
polyprotein (monocistronic RNA construct), and
(B) within the 3' NTR (bicistronic RNA
construct).

10 **Fig. 6:** Diagram of the structure of HCV-RNA constructs
with inserted reporter genes (A) as part of a HCV
replicon from NS3 to NS5B; - in the end the
reporter protein is cleaved by viral or cellular
proteases out of the polyprotein or the marker
gene for selection (selection gene) or the
resistance gene are transfected into the cells
through co-transfection -, (B) as part of a
fusion gene composed of a resistance and reporter
gene (for example for the neomycin
phosphotransferase and green fluorescent
protein), (C) as part of a replicon composed of a
resistance and reporter gene (for example the
neomycin phosphotransferase and green fluorescent
protein) bound via a nucleotide sequence, which
encodes for an amino acid sequence (area with
hatches), which can be cleaved by a protease or
has a self-cleaving (autocatalytic) activity, (D)
as independent gene (in this case the green
fluorescent protein), which is expressed from its
own internal ribosome binding site (IRES); - the
resistance gene (in this case the neomycin
phosphotransferase gene) is also expressed from
its own internal ribosome binding site (IRES)
(polycistronic construct).

35 **Fig. 7:** Diagram of the structure of a HCV-RNA construct,
where the resistance gene is linked to the HCV-

RNA sequence via a ribozyme or a recognition site for a ribozyme.

The thick lines illustrate the HCV 5' and 3' NTRs, E-I is a heterologous internal ribosome binding site, which is required for the expression of the resistance gene, and the grey square illustrates the ribozyme or a recognition site for a ribozyme.

Fig. 8: Diagram of the structure of a HCV-RNA construct with resistance gene and integrated foreign gene.

Fig. 9: Method for comparing the specific infectivity (expressed as number of cell colonies obtained) of total RNA against in vitro transcripts. HCV-RNA is generated by in vitro transcription of a corresponding RNA construct and quantified by measurement of the optical density at 260 nm (OD 260 nm). A defined number of these molecules is mixed with a specified amount of total RNA from naive Huh-7 cells and this mixture is transfected into naive Huh-7 cells with the aid of electroporation. At the same time the total RNA of a cell clone, produced by the method described in Figure 1, is isolated using a known state of the art method and the amount of HCV-RNA contained therein is determined by means of Northern Blot using a HCV specific RNA probe and subsequent quantification via phosphoimaging. A defined amount of this total RNA is analogously transfected in naïve Huh-7 cells. These cells from both the cultures are then subjected to a G418 selection and the number of colonies created is determined by counting after fixing and staining with Coomassie Brilliant Blue. For the determination of transfection efficiency 1 μ g of a

plasmid allowing the expression of luciferase is added to each transfection culture. An aliquot of the transfected cells is collected after 24 hours and the luciferase activity determined in the respective cell lysates. The number of colonies is always normed to the luciferase expression.

Fig. 10: Sequence analysis of the 9-13 clone. Total RNA of the 9-13 cell clone, resulting from transfection of HCV-RNA construct I377/NS3-3', was isolated using a known state of the art method and the HCV-RNA construct amplified from nucleotide position 59 to 9386 with the aid of 'long-distance RT-PCR' using primer S59 and A9413. The PCR fragments were cloned and 11 clones (9-13 A - K) completely sequenced. Clones D and I, E and G as well as H and J turned out to be identical, respectively. The positions of the amino acid differences in the NS3-5B region between the recloned HCV-RNAs and the parental construct are marked with a thick vertical mark for each clone. Each clone was digested with restriction enzyme *Sfi*I and the respective fragment inserted in the parental construct. These clones were each transfected in Huh-7 cells and the cells subjected to selection as described in Figure 1. The number of cell clones obtained with each construct is noted next to the respective construct on the right.

Fig. 11 A: Principle of determination of replication with the aid of a reporter gene. In the upper part of the figure, the HCV-DNA construct I₃₈₉/Luc/NS3-3' is shown. It consists of the HCV 5' NTR (nucleotide positions 1-389), the luciferase gene (*luc*), the IRES of the encephalomyocarditis virus, the HCV NS3-5B and the 3' NTR. The

position of the active centre of the NS5B RNA polymerase, into which a deactivating amino acid exchange was introduced, is indicated by 'GND'. The plasmids, which code for the HCV-^{RNA} construct which is able to replicate or is defective, are digested with the restriction enzyme *Sca*I and added to an in vitro transcription reaction with T7 RNA polymerase. After removal of the template DNA, the respective HCV-^{RNA} constructs were transfected in naive Huh-7 cells by means of electroporation and the latter collected at regular intervals.

Fig.11 B: Comparison of luciferase activity in cells transfected with the parental HCV-^{RNA} construct I₃₈₉/Luc/NS3-3'/wt (wt) or the following variants: inactive RNA (318 DN), variants 9-13F or variant 5.1. The cells were collected 6 (not shown), 24, 48, 72, 96, 120, 144 and 168 hours after transfection and luciferase activities determined by luminometric measurement.

Fig. 12: Selectable HCV full-length genomes (constructs I₃₈₉/core 3'/5.1 and I₃₈₉/core 3'/9-13F).
25 (A) Diagram of the full-length construct. The region between both indicated recognition sites for the restriction enzyme *Sfi*I corresponds to the sequences of the highly adapted RNA variants 5.1. or 9-13F.
30 (B) Number of colonies which were obtained after transfection of 0.1 µg in vitro transcribed RNA of the construct I₃₈₉/core-3'/5.1. described under A into HUH7-cells. The result of a representative experiment is given.
35 (C) Demonstration of autonomously replicating HCV full-length RNAs in G418 resistant cell clones which were obtained after transfection of the

corresponding in vitro transcript. The illustration shows the autoradiogram of a Northern Blot, which was hybridised with a probe against the neo-resistance gene and the HCV 5' NTR. The controls shown in lanes 1 and 2 each correspond to 10^8 molecules of the indicated in vitro transcripts, mixed with total RNA from naive Huh-7 cells. The negative control contains only total RNA from naive Huh-7 cells (lane 3). Lanes 4-9 contain 3-10 µg total RNA from G418 resistant cell clones, which were obtained after transfection by in vitro transcribed I₃₈₉/core 3'/5.1 RNA or I₃₈₉/core 3'/9-13F RNA. The G418 concentration used for the selection is indicated in each case. Five of the cell clones shown contain the highly adapted RNA variant 5.1 (lanes 4-8), one contains the adapted RNA variant 9-13F (lane 9).

20 Fig. 13: HCV-RNA constructs with a reporter gene. (A) Bicistronic HCV-RNA constructs. The reporter gene is translated with the aid of a separate IRES. (B) Monocistronic HCV-RNA constructs. The reporter gene product is expressed as fusion protein with a HCV protein. Both portions are linked via a recognition sequence for a viral or cellular protease, which permits a proteolytic separation of the two fused protein portions. In the example shown the reporter gene product and respective HCV protein was fused through a recognition sequence for ubiquitin (Ub).

25 30 35 Fig. 14: Tricistronic full-length HCV-RNA construct, that in addition to the resistance gene possesses an inserted foreign gene (ther. gene).

Fig. 15: Monocistronic HCV-RNA constructs, for which the resistance gene product is expressed as a fusion protein with HCV portion. The resistance gene (RG) is either active as a fusion protein or it is fused with the HCV portion via a proteolytically cleavable sequence in such a way that the resistance gene product is split from the HCV portion by a cellular or viral protease. In the example shown the resistance gene was fused with the respective HCV portion through the sequence coding for ubiquitin (Ub).

15 Example 1: Production of HCV-RNA constructs

(A) Synthesis and cloning of a complete HCV consensus genome by means of RT-PCR

20 The HCV genome, which means the HCV-RNA, was isolated from
the liver of a chronically infected patient as described in
the following:

The total RNA was isolated from approx. 100 mg of liver according to the method described by Chomczynski and Sacchi (1987, Anal. Biochem. 162, 156). Using 1 µg of this isolated RNA a reverse transcription was carried out with the primers A6103 (GCTATCAGCCGGTTCATCCACTGC) or A9413 (CAGGATGGCCTATTGG CCTGGAG) and the expand reverse transcriptase system (Boehringer Mannheim, Germany) according to the manufacturer's recommendations. A polymerase chain reaction (PCR) was carried out with the products of this reverse transcription using the expand long template system (Boehringer Mannheim, Germany), in which the buffer containing 2 % of diméthyl sulfoxide was used. After one hour at 42 °C, 1/8 of the reaction mixture was used for a first PCR with primers A6103 and S59.

(TGTCTTCACGCAGAAAGCGTCTAG) or A9413 and S4542 (GATGAGCT CGCCCGCGAAGCTGTCC). After 40 cycles, 1/10 of this reaction mixture was used for a second PCR with primers S59 and A4919 (AGCACAGCCCCGCGTCATAGCACTCG) or S4542 and A9386 (TTAGCTCCCCG TTCATCGGTTGG. After 30 cycles the PCR products were purified by means of preparative agarose gel electrophoresis and the eluted fragments were ligated into the vector pCR2.1 (Invitrogen) or pBSK II (Stratagene). Four clones of each fragment were analyzed and sequenced, and a consensus sequence was established. For this purpose the DNA sequences were compared to each other. The positions, where the sequence of one fragment was different from the others, were considered as undesired mutations. In the case of ambiguities of the sequence, shorter overlapping PCR fragments of the respective region were amplified and several clones sequenced. By this means several potential mutations could be identified in each fragment and consequently an isolate specific consensus sequence could be established. This established consensus sequence or genome belongs to the worldwide spread genotype 1b. The non-translated region at the 3' end (=3' NTR) was obtained by means of a conventional PCR, whereby an antisense primer, which covers the last 24 nucleotides of the 'X-tail' known from the state of the art was used (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 744; und Rice, PCT/US 96/14033). The authentic non-translated region on the 5' end (=5' NTR) downstream of the T7 promoter was created by means of PCR, whereby an oligonucleotide, which corresponds to a shortened T7 promoter (TAA TAC GAC TCA CTA TAG) and the first 88 nucleotides of HCV, was used on one hand and one of the previously mentioned plasmids, which carries one of the 5' fragments of the genome, was used on the other hand. A complete HCV consensus genome was assembled from the subgenomic fragments with the smallest number of non-consensus replacements and inserted into a modified pBR322 vector. Deviations from the consensus sequence were

eliminated by means of site-directed mutagenesis. In order to produce run-off transcripts with an authentic 3' end, the 3' NTR of the isolates (with the end TGT) was modified to AGT (according to the sequence of the genotype 3' = clone 5 'WS' according to Kolykhlov et al., 1996, J. Virol. 70, 3363) and an additional nucleotide replacement was carried out at position 9562, to preserve the A:T pairing in the hairpin structure at the 3' end of the 3' NTR (Kolyhalov et al. ibid.). In order to eliminate an internal restriction 10 site for the Scal enzyme, a silent nucleotide replacement was further carried out. After joining the full-length genome with the appropriate 5' and 3' NTRs the complete HCV sequence was analyzed. No undesired nucleotide replacement was detected.

15

The HCV genome produced in this way should be hepatotropic according to the definition.

20

(B) Synthesis of selectable HCV subgenomic constructs

By means of the consensus genome described under (A), HCV subgenomic constructs were created, which include the antibiotic resistance gene neomycin phosphotransferase (NPT) and two sequences of internal ribosome entry sites (IRES). The biochemical procedures used for this are known and familiar to the expert (see: Sambrook, J., E.F. Fritsch, T. Maniatis, 1989, Molecularcloning: a laboratory manual, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbor, N.Y.; Ausubel et al. (eds.), 1994, Current Protocols in Molecular Biology, Vol. 1-3, John Wiley & Sons Inc., New York). The antibiotic resistance gene was inserted immediately after the 5' NTR, obtaining a bicistronic RNA as a result (see Fig. 1 A). However, the antibiotic resistance gene might as well be inserted at another site of the HCV subgenomic construct, for example within the nucleotide sequence encoding for the polyprotein, obtaining a monocistronic RNA as a result (see

Fig. 5 A), or in the 3' NTR (see Fig. 5 B). The IRES elements are one of the two HCV-IRES variants from nucleotides 1-377 or nucleotides 1-389 as well as the IRES of the encephalomyocarditis virus, which controls the 5 translation of the HCV sequence downstream from the genes for NS2 or NS3 up to the authentic 3' end of the genome.

The two HCV-IRES variants mentioned were determined in the following way:

10

Based on deletion analyses of the 3' end of the HCV-IRES (Reynolds et al. 1995, EMBO J. 14, 6010) various segments of the 5' NTR were fused with the NPT gene and analyzed with regard to the maximum number of colonies formed by means of 15 co-transfection with a plasmid containing the T7 RNA polymerase gene. The best results were achieved with the HCV sequences from 1-377 to 1-389. Due to the fact that the AUG start codon of the HCV polyprotein is located at position 342 and is consequently included in the IRES sequence, 12 or 16 amino acids of the HCV core protein ("core protein") were fused with the neomycin phosphotransferase (see Fig. 1 A).

Accordingly, these modified HCV subgenomic constructs were 25 given the designations $I_{377}/NS2-3'$ (or $I_{377}/NS3-3'$) and $I_{389}/NS2-3'$ (or $I_{389}/NS3-3'$). They are illustrated in a diagram in Fig. 1A.

Different cell lines and primary cell cultures of human 30 hepatocytes were transfected with in vitro transcripts of these modified parental HCV subgenomic constructs $I_{377}/NS2-3'$ (or $I_{377}/NS3-3'$) and $I_{389}/NS2-3'$ (or $I_{389}/NS3-3'$).

As a parallel negative control for all transfection 35 experiments, an appropriately modified, but defective subgenome was constructed for every modified parental HCV subgenomic construct, which differs from the parental

construct due to the fact, that it has a deletion of 10 amino acids within the reading frame comprising the active site of the NS5B RNA polymerase (Behrens et al., 1996, *EMBO J.* 15, 12; and Lohmann et al., 1997, *J. Virol.* 71, 8416).

5 (C) Synthesis of selectable HCV genomic constructs

A NS2-3' subgenomic construct, linked at its 5' end to a fragment of the luciferase gene and the complete EMCV-IRES, 10 was restricted with NcoI and SpeI and purified using preparative agarose gel electrophoresis. The vector obtained in this way was ligated with a NcoI/NotI-HCV fragment corresponding to the nucleotide positions 342 to 1968 of the HCV genome and a NotI/SpeI fragment 15 corresponding to the nucleotide positions 1968-9605 in a 3-factor-ligation. The resulting construct, where the complete HCV open reading frame and the 3' NTR lie downstream from the luciferase gene fragment and the EMCV-IRES, was then restricted with PmeI and SpeI and ligated 20 with the analogously restricted I₃₈₉/NS3-3'/wt subgenomic construct vector. This selectable HCV genomic construct is illustrated in Fig. 4.

25 (D) Production of in vitro transcripts corresponding to the HCV-RNA constructs

The previously described purified plasmid DNAs were linearized with ScaI and used for an in vitro transcription reaction after phenol/chloroform extraction and isopropanol precipitation using the following components: 80 mM HEPES, 30 pH 7.5, 12.5 mM MgCl₂, 2 mM Spermidine, 40 mM Dithiothreitol, 2 mM of each NTP, 1 unit RNasin/ μ l, 50 μ g/ml restricted DNA and approx. 2 units/ μ l T7 RNA polymerase. After 2 hrs. at 37 °C half of the amount of T7 35 polymerase was added and the reaction mixture was incubated for two further hours. In order to remove DNA the mixture was extracted with acid phenole (U. Kedzierski, J.C. Porte,

1991, Bio Techniques 10, 210), precipitated with isopropanol, the pellet was dissolved in water and incubated with DNase (2 units per μ g DNA) for 60 min. at 37 °C. After subsequent extraction with acid phenole, acid phenole/chloroform and chloroform as well as isopropanol precipitation the dissolved RNA was quantified using optical density measurement and its integrity was checked using formaldehyde-agarose gel electrophoresis.

10

Example 2: Transfection experiments with the hepatoma cell line Huh 7

With all transfection experiments it was carefully ensured
15 that any template DNA had been removed beforehand so as to avoid the possibility of this DNA integrating in transfected cells and conferring a neomycin resistance to them independent from HCV replication. The reaction mixture was therefore treated with 2 units of DNase per μ g DNA for 60 min. at 37 °C after the in vitro transcription and extracted with acid phenole, acid phenole/chloroform and chloroform. Prior to being used for the transfection the precipitated RNA was analyzed using formaldehyde-agarose gel electrophoresis.

25

Three separate transfection experiments were carried out with the highly differentiated human hepatoma cell line Huh 7 (according to Nakabayashi et al. 1982, Cancer Res. 42, 3858). Each time 15 μ g RNA were introduced into 8×10^6 Huh 30 7 cells by electroporation and the cells were seeded in culture dishes with a diameter of 10 cm. 24 hours after seeding, neomycin (= G418) was added in a final concentration of 1 mg/ml. The culture medium was changed twice a week. After 3 - 5 weeks small colonies were 35 visible, which were isolated and grown under the same culture conditions.

The cell clones obtained during the first experiment were isolated and subpassaged. Most of the clones died during this procedure, and the final yield was only 9 clones from cells, which had been transfected with the parental HCV subgenomic constructs and 1 clone (clone 8-1) from cells, which had been transfected with a defective HCV genomic construct, namely a defective NS2-3' HCV-RNA. Apart from an extended doubling time and the occasional occurrence of irregularly shaped cells, no consistent morphological differences were found between these 9 cell clones and the single cell clone (clone 8-1) or the parental Huh 7 cells.

The main criteria for functioning HCV genomic constructs are the formation of viral RNA with the correct size and the absence of (integrated) plasmid DNA, which could transfer or mediate on a G418 resistance.

To determine the HCV-RNA in the Huh 7 cells, the total RNA was isolated and analyzed by means of the common Northern Blot method using a plus-strand specific ribo probe (RNA probe). For this purpose the total RNA was isolated from the respective cell clones according to the method described by Chomczynski and Sacchi 1987, *Anal. Biochem.* 162, 156 and 10 µg RNA, which is equivalent to the total RNA content of 0,5 - 1 x 10⁶ cells, are separated using denatured formaldehyde-agarose gel electrophoresis (lanes 3 to 12 of Fig. 1 B). At the same time 10⁹ in vitro transcripts (ivtr.), which correspond to the I₃₈₉/NS2-3'/wt or the I₃₈₉/NS3-3'/wt replicon RNAs, are separated as well as size markers with authentic sequence (lane 1 or lane 2). The separated RNA was transferred onto nylon-based membranes and hybridized with a radioactively labeled plus-strand specific RNA probe, which was complementary to the complete NPT gene and the HCV-IRES of nucleotide 377 to nucleotide 1. The positions of the HCV specific RNAs (arrows) and the 28S rRNA are specified to the right of lane 12, the size (amount of nucleotides) of the RNA marker

is specified to left of lane 1. The RNA marker fragments contain HCV sequences and therefore hybridize with the ribo probe (= RNA probe). The results of this analysis are illustrated in Fig. 1 B.

With the exception of clone 8-1 transfected with the defective HCV genomic construct, all cell clones produced homogenous HCV-RNAs of correct lengths (approx. 8640 nucleotides in the case of the NS2-3' and approx. 7970 nucleotides in the case of NS3-3' replicon). This result is an indication of the fact, that the functional replicons or the functional HCV genomic constructs transfer the G418 resistance. In order to exclude, that the G418 resistance is caused by a plasmid DNA, which is integrated in the genome of the Huh 7 host cell and transcribed under control of a cellular promoter, the DNA of each clone was analyzed by means of a NPT gene specific PCR. Consequently the DNA was isolated from the selected Huh 7 cell clones by means of digestion with proteinase K (40µg/ml, 1h, 37 °C) in 10mMTris, pH7,5, 1mM EDTA, 0,5 % SDS and subsequent extraction with phenol, phenol/chloroform and isopropanol precipitation. The DNA precipitate was dissolved in 10 mM Tris (pH 7,5) and 1 mM EDTA and incubated with Rnase A for 1 hour. Following a phenol/chloroform extraction and ethanol precipitation, 1 µg DNA, equivalent to 4 - 8 x 10⁴ cells, was analyzed by means of PCR using NPT gene specific primers (5'-TCAAGACCGACCTG TCCGGTGCCC-3' and 5'- CTTGAGCCTGGCGAACAGTTGGC-3'), and a DNA fragment consisting of 379 nucleotides was generated. The specificity of the PCR product was established by means of the Southern Blot method, in which a DNA fragment labeled by digoxigenin was used, which corresponds to the NPT gene. As positive controls (for the detection of possibly contaminating nucleic acids) the PCR method was carried out with 10⁷ plasmid molecules or 1 µg DNA from a BHK cell line, which was stably transfected with a neomycin resistance gene, and

as negative control, the PCR was carried out with the same reagents but without added DNA.

- The results of this analysis are illustrated in Fig. 1 C.
5 The lanes 1 and 2 represent the positive controls, lane 13 represents the negative control. The numbers to the left of lane 1 indicate the sizes of the nucleotide marker molecules.
- 10 A NPT-DNA could not be detected in any cell clone, apart from clone 7-3 (Fig.1C, lane 3), which was obtained from cells after transfection with a NS2-3' replicon/NS2-3'HCV genomic construct, and clone 8-1 (Fig.1C, lane 12), which was obtained from cells after transfection with a defective HCV genomic construct. This result was another indication of the fact, that the G418 resistance of most clones was passed on by the replicating HCV-RNA. But even regardless of these results, it is unlikely, that integrated plasmid DNA produces HCV-RNAs of correct sizes, because the 15 plasmids used for the in vitro transcription contain neither an eukaryotic promoter nor a polyadenylation signal. In the case of clone 7-3 the resistance is therefore very probably passed on by the HCV-RNA construct or the replicating HCV-RNA as well as by an integrated NPT 20 DNA sequence, whereas the resistance of the cells of clone 8-1 is only caused by the integrated plasmid DNA.

- Clone 9-13 (Fig. 1 B, lane 11) was subject to further tests, to confirm that the G418 resistance was passed on by 30 a self-replicating HCV-RNA. Clone 8-1, which carries the integrated copies of the NPT gene, was used throughout as negative control. A PCR was carried out, which allows the detection of < 1000 NPT gene copies in ~ 40.000 cells, with the aim to rigorously exclude the presence of NPT-DNA in 35 clone 9-13. The result of this PCR is illustrated in Fig. 2A. The PCR proceeded in detail as follows:

During the test, 10^6 - 10^2 plasmid molecules ($I_{377}/NS3-3'$ /wt) were used either directly (lanes 7 - 11) or after adding each 1 μ g 9-13 DNA (lanes 2 - 6). The specificity of the amplified DNA fragment was determined by Southern Blot using a NPT specific probe. A PCR without DNA probe was carried out as negative control (lane 12).

Even with this sensitive method, no plasmid DNA could be detected in one μ g DNA of the cell clone 9-13 (lane 1). To estimate the amount of HCV plus- and minus-strand RNAs in these cells, a dilution series of total RNA was analyzed with the Northern Blot method using a plus- or minus-strand specific radioactively labeled ribo probe (= RNA probe). For this purpose, 8, 4 or 2 μ g of total RNA, which have been isolated from the cell clones 9-13 and 8-1, were analyzed in parallel to known amounts of in vitro transcripts with plus- or minus-strand polarity (control RNAs) in the Northern Blot method and were then subjected to a hybridization. The hybridization was carried out with a plus-strand specific ribo probe, which covered the complete NPT gene and the HCV-IRES ('plus-strand', top panel), or with a minus-strand specific RNA probe, which was complementary to the NS3 sequence ('minus-strand', bottom panel). The arrows mark the positions of replicon RNA. The results of this analysis are illustrated in Fig. 2.B.

Approx. 10^8 copies/ μ g total RNA were detected in the case of the plus-strand, which is equivalent to 1000 - 5000 HCV-RNA molecules per cell, whereas the amount of minus-strand RNA was 5- to 10-times less. This result corresponds to the assumption, that the minus-strand RNA is the replicative intermediate form or intermediate copy, which is used as a template for the synthesis of the plus-strand molecules.

35

Due to the fact that the reaction is mainly catalyzed by the viral RNA-dependent RNA polymerase, the synthesis of

the HCV-RNAs should be resistant to dactinomycin, an antibiotic, which selectively inhibits the RNA synthesis of DNA templates, but not the RNA synthesis from RNA templates. To confirm this assumption, cells were incubated with [³H] uridine in the presence of dactinomycin, the radioactively labeled RNAs were extracted, separated by means of denaturing agarose gel electrophoresis and analyzed with the aid of a common Bio-Imager using a [³H]-sensitive screen. For this purpose approx. 5 x 10⁵ cells of the clones 9-13 and 8-1 were incubated at a time with 100 µCi [³H] uridine for 16 hrs. in the absence (-) or presence (+) of 4 µg/ml of dactinomycin (Dact). Following this labeling reaction the total RNA was prepared and analyzed by means of formaldehyde-agarose gel electrophoresis. Only 1/10 of the total RNA is illustrated in the first two lanes. The radioactively labeled RNA was visualized using a BAS-2500 Bio-Imager (Fuji).

The results of this analysis are illustrated in Fig. 2 C. In accordance with the inhibitor profile of the NS5B polymerase (Behrens et al., 1996, EMBOJ. 15, 12 and Lohmann et al., 1997, J. Virol. 71, 8416), dactinomycin had no influence on the replication of the HCV RNA, whereas the synthesis of cellular RNA was inhibited. A RT-PCR was carried out for recloning the replicating sequences, to confirm the identity of the viral RNA. The sequence analysis of the recloned RNA showed that the RNA in clone 9-13 is HCV specific and corresponds to the transfected transcript of the HCV construct I₃₇₇/NS3-3'/wt.

For the analysis of the viral proteins, at first the respective cells were metabolically radioactively labeled with [³⁵S] methionine/cysteine, then lysed and afterwards the HCV specific proteins were isolated from the cell lysates by means of immunoprecipitation. The results of these analyses are illustrated in Fig. 3 A. The detailed procedure was as follows: Cells of the cell clones 9-13

(wt) and 8-1 (Δ) were metabolically radioactively labeled by treating them with a protein labeling mixture familiar to the expert and available on the market (for example. NEN Life Science). The HCV specific proteins were separated from the cell lysate by immunoprecipitation (IP) under non-denaturing conditions (for example according to Bartenschlager et al., 1995, J. Virol. 69, 7519) using three different antisera (3/4, 5A, 5B, according to the labeling on the top end of the lanes 1 to 12). The immune complexes were analyzed by means of tricine SDS-PAGE and made visible by means of autoradiography. To obtain authentic size markers, the homologous replicon construct I_{3,77}/NS3-3'/wt was subject to a transient expression by the vaccinia virus-T7 hybrid system in the Huh 7 cells. The resulting products were used as size markers (lanes 7 - 9) parallel to the cells of the clones 9-13 and 8-1. Identified HCV proteins are labeled on the left edge of lane 1, the molecular weights (in Kilodalton) are specified on the right edge of lane 9. It should be noted that the NS3/4 specific antiserum ('3/4') used preferably reacts with NS4A and NS4B causing an underrepresentation of NS3.

All viral antigens could unambiguously be detected, and their apparent molecular weights did not show any difference to those being detected after transient expression of the same bicistronic HCV-RNA construct in the original Huh 7 cells. An immunofluorescence detection reaction was carried out using NS3 and NS5A specific antisera, to determine the subcellular distribution of the viral antigens (for example according to Bartenschlager et al., 1995, J. Virol. 69, 7519). For this purpose cells of the clones 9-13 (wt) and 8-1 (Δ) were fixed with methanol/acetone 24 hrs. after incubating on coverslips and incubated with polyclonal NS3 or NS5A specific antisera. The bound antibodies were made visible with a commercially available FITC conjugated anti-rabbit antiserum. The cells

were counterstained with the 'Evans Blue' stain to suppress unspecific fluorescence signals.

The results of this detection test are illustrated in Fig. 5 3 B. A strong fluorescence in the cytoplasm could be detected with both antisera. The NS5A specific antiserum also caused a slight nuclear fluorescence, which indicates that at least small amounts of this antigen also reach the nucleus. But the generally dominating presence of the viral 10 antigens in the cytoplasm are a strong indication that HCV-RNA replication occurs in the cytoplasm, as is the case with most RNA viruses.

These results prove clearly that the establishment of a 15 cell culture system for the HCV could be accomplished with the test arrangement described, the efficiency of which surpasses everything known up until now by far and for the first time allows the detection of viral nucleic acids and proteins with conventional and approved biochemical 20 methods. This efficiency actually allows detailed examination of HCV pathogenesis, genetic analyses of different HCV functions and a precise study of the virus/host cell interaction, through which new starting points for the development of a antiviral therapy can be 25 defined.

Example 3: Transfection of Huh 7 cells with HCV total constructs

30

Huh7 cells are transfected and selected as described in example 2, whereby in this case selectable constructs are used, which contain the complete virus genome. Corresponding to Example 2, the resulting cell clones are 35 tested for the absence of HCV-DNA by means of PCR and the productive replication of HCV-RNA is then established by means of Northern Blot, [³H]uridine labeling in the

presence of dactinomycin, detection of the viral proteins or antigens preferably with the aid of the Western Blot, the immunoprecipitation or immunofluorescence. In contrast to the arrangements described in Example 2, the construct described here makes it possible to obtain more complete and very likely infectious viruses, which has not been the case in the subgenomic constructs described in Example 2. These viruses existing in the cell and the cell culture supernatant are concentrated for example by means of ultracentrifugation, immunoprecipitation or precipitation with polyethyleneglycol, and all exogenous nucleic acids, which means those that are not incorporated into the virus particle, are digested by incubation with nucleases (RNase, DNase, micrococcal nuclease). In this way all contaminating nucleic acids, which are not included in the protecting virus particle, can be removed. After inactivation of the nucleases, the protected viral RNA is isolated for example by means of incubation with proteinase K in a buffer containing SDS, by extracting with phenol and phenol/chloroform and detected by means of Northern Blot or RT-PCR using HCV specific primers. Also in this test arrangement, the combination of the HCV consensus genome described with a selection marker was crucial for the efficient production of viral RNA, viral protein and therefore HCV particles.

Example 4: Production and application of a HCV-RNA construct, whereby the resistance gene is linked to the HCV subgenomic sequence via a ribozyme or a recognition site for a ribozyme.

A HCV-RNA construct is produced according to Example 1 or Example 3, where an antibiotic resistance gene is linked to the HCV-RNA sequence through a ribozyme or a recognition sequence for a ribozyme. These constructs are illustrated in a diagram in Fig. 7. Huh 7 cells are transfected with

this HCV-RNA construct as described in Example 2. A selection with the appropriate antibiotic follows the transfection into the cells. The inserted ribozyme is activated in the cell clones obtained in the procedure or, in the case of a construct, which carries a recognition sequence for a ribozyme, the ribozyme is transfected into the cell (for example by means of transfection of a ribozyme construct or infection with a viral expression vector, into which the respective ribozyme has been inserted). In both cases the resistance gene is separated from the HCV-RNA sequence by the ribozyme-dependent cleavage. The result in the case of the HCV genome is an authentic HCV genome without a resistance gene, which can form authentic infectious virus particles. A HCV replicon without resistance gene is created in the case of the HCV subgenomic constructs.

Example 5: Co-transfection of a HCV-RNA construct with a separate luciferase transfection construct

A HCV-RNA construct is produced according to Example 1 (A) or Example 3 or Example 4. At the same time a transfection construct is produced, which comprises the luciferase gene, whereby this luciferase gene is linked to a first nucleotide sequence, which encodes a HCV protease (for example NS3 protease) cleavage site, to a second nucleotide sequence, which encodes for another protein or a part of another protein. HCV-RNA construct and transfection construct are transfected into any host cells, preferably hepatoma cells, most preferably Huh 7 cells. This can be realised as described in Example 2. The product of the modified luciferase gene is a luciferase fusion protein, where the luciferase is inactivated due to the fusion with the foreign part. The fusion protein, which contains a recognition sequence for a HCV protease, is cleaved in transfected cells with high HCV replication, and

consequently the active form of the luciferase, which can be identified through luminometric measurement, is released. If the replication of the HCV-RNA construct is inhibited the fusion protein will not be cleaved and no active luciferase will be released. The quantitative determination of the luciferase is therefore a measure for the replication of the HCV construct. Instead of the luciferase gene, another reporter gene can just as easily be used, which is modified in the same way, so that its expression depends on the viral replication, although this reporter gene is not part of the HCV construct. A cellular protein, which is deactivated or activated by the HCV proteins or nucleic acid, can also be used as a so called surrogate marker. The expression or activity of this surrogate marker is in this case a measure for the replication of the viral DNA.

20 Example 6: Production of HCV subgenomic constructs with integrated foreign genes to be used as liver cell specific vector in gene therapy

These recombinant and selectable HCV subgenomic constructs are transfected in transcomplementing helper cell lines, 25 which means in cell lines, which express the missing functions (for example the structural proteins) in an inducible or constitutive way. Cell clones containing a functional HCV subgenomic construct can be established through appropriate selection. The viral structural 30 proteins expressed from the host cell allow the formation of virus particles, into which the RNA of the HCV subgenomic constructs will be transfected. The results are therefore virus-like particles, which contain a HCV subgenomic construct according to the invention including 35 the inserted foreign gene and which can transmit this to other cells by means of infection. An example for this construct is illustrated in Fig. 8.

It is also possible to use the HCV subgenomic construct with integrated foreign gene directly as an expression vector. This involves the same method as mentioned previously except that cell lines, which do not express transcomplementing factors, are transfected. In this case the HCV construct is only used as an expression vector.

10 **Example 7: Production of cell culture adapted HCV-RNA constructs**

(A) **Method of isolation**

15 The following method was used to determine adaptive mutations and to produce cell culture adapted HCV-RNA constructs: cells were transfected with a HCV-RNA construct as described in Examples 1 and 2 and G418-resistant cell clone produced. For the determination of ability to
20 replicate (understood in this context to be the number of G418 resistant cell clones obtained per microgram of transfected HCV-RNA or HCV-RNA construct), the total RNA from one of the cell clones, [the 9-13 clone (Fig. 1 B, lane 11)], was isolated and the quantity of HCV-RNA contained within it was determined by Northern Blot as described in Fig. 2 B. Ten micrograms of the total RNA, containing approx. 10^9 molecules of HCV-RNA, was then transfected into naive Huh-7 cells using electroporation (Fig. 9). In parallel, 10^9 in vitro transcripts of the
25 analogous neo-HCV-RNA, which had been adjusted with isolated total RNA from naive Huh-7 cells to a total RNA quantity of 10 μ g, were transfected in naive Huh-7 cells. After selection with G418, the number of cell colonies, expressed as 'colony forming units (cfu) per microgram
30 RNA', was determined in both these cultures. At a concentration of 500 μ g/ml G418 in the selection medium, the number of colonies obtained with the HCV-RNA contained
35

in isolated total RNA from clone 9-13, was approx. 100,000 cfu per microgram HCV-RNA. In contrast, only 30 - 50 colonies were obtained with the same quantity of in vitro transcribed HCV-RNA. This result confirms that the specific infectivity of the HCV-RNA isolated from the cell clones is approx. 1,000 - 10,000 times higher than the infectivity of the analogous in vitro transcripts. The experimental approach is shown in Fig. 9.

With the aid of 'long-distance RT-PCR', the HCV-RNA was amplified from the total RNA of the 9-13 cells, the PCR amplificate was cloned and numerous clones were sequenced. A comparison of the sequences of these recloned RNAs with the sequence of the RNA originally transfected into the naive Huh-7 cells, showed that the recloned RNAs possessed numerous amino acid exchanges distributed over the whole HCV sequence (Fig. 10). *SfiI* fragments of these recloned mutants were used to replace the analogous *SfiI* fragment of the original replicon construct, and RNAs of the respective mutants were transfected in naive Huh-7 cells. After selection with G418 the number of colonies created was determined for each HCV-RNA mutant. While only 30 - 50 colonies per microgram RNA were obtained with the parental RNA the number of colonies was noticeably higher for two of the recloned variants (Fig. 10). In the case of the HCV-RNA constructs 9-13I and 9-13C the specific infectivity was increased to 100 - 1,000 cfu per microgram RNA and for 9-13F replicon it was 1,000 - 10,000 cfu per microgram RNA. These results show that the amino acid exchanges in the analysed NS3-5B regions of the mutants 9-13I, 9-13C and particularly of 9-13F, led to a considerable increase in ability to replicate. In contrast all the other HCV-RNA constructs (9-13 A, B, G, H und K) were no longer able to replicate; they thus contained lethal mutations.

35

In order to answer the question which of the amino acid exchanges in the 9-13F construct led to an increase in

replication, the exchanges were introduced separately or in combination into the parental HCV-RNA construct, and the corresponding RNAs transfected in naïve Huh-7 cells. The result of the transfection with these RNAs is summarised in Table 1. From this it is evident that in the present example the high ability to replicate is determined by several mutations. The amino acid exchanges in the HCV-RNA regions NS5A and NS4B make the greatest contribution. The single exchanges in the NS3-Region also make a contribution and perhaps they are synergistic.

These results confirm that it was through the G418 selection of the cells transfected with the neo-HCV-RNA construct that there was enrichment of those HCV-RNAs having noticeably higher ability to replicate. HCV-RNA constructs with greatly differing replication efficiencies can be selected using the experimental approach described here. The higher the concentration of the antibiotic in the selection medium, in/on which the HCV-RNA construct containing cells are cultivated for selection, the higher must be the extent of adaptive mutations and hence replication efficiency of the relevant HCV-RNA constructs, to allow the cells to grow under these conditions. If the selections are carried out using lower antibiotic concentrations, cells can survive and multiply, but the HCV-RNA construct shows a comparatively lower replication efficiency and fewer adaptive mutations.

As has been shown, the 9-13F HCV-RNA construct described so far, which contains several adaptive mutations, had a higher replication efficiency than the parental HCV-RNA. In order to obtain HCV-RNAs with even higher replication in cell culture, the HCV-RNA contained in the total RNA of a selected cell clone was passaged several times in naïve Huh-7 cells. The selected 5-15 cell clone, was obtained by transfection with the HCV-RNA construct I₃₈₉/NS3-3' (Fig. 1). It largely corresponds to the cell clone 9-13, produced

by transfection with a HCV-RNA construct, having a HCV-IRES shorter by 22 nucleotides ($I_{377}/NS3-3'$; Fig. 1). Ten micrograms of total RNA, isolated from cell clone 5-15, were transfected into naïve Huh-7 cells using electroporation and the cells subjected to a selection with 1 mg/ml G418. The total RNA from one of the cell clones thus produced was again isolated, transfected into naïve Huh-7 cells and selected in the same way. This process was repeated a total of four times. After the fourth passage 10 the total RNA was isolated from a cell clone and the neo-HCV-RNA amplified with the aid of the 'long-distance RT-PCR'. The amplified DNA fragment was digested with the restriction enzyme *Sfi*I and inserted into the *Sfi*I-restricted parental construct $I_{389}/NS3-3'$. Over 100 DNA 15 clones were obtained altogether and then analysed by means of restriction digestion. In vitro transcribed RNA of about 80 of these clones was each transfected into naïve Huh-7 cells and subjected to a selection with 500mg/ml G418. Of the 80 neo-HCV-RNA variants examined, the great majority 20 proved to be replication defective. However, the specific infectivity, expressed as 'colony forming units' per microgram RNA, was noticeably increased in the case of two mutants, 5.1 and 19 (Table 2). Through several passages of the RNA in cell culture it is clear that HCV-RNAs are 25 produced whose replication efficiency due to mutations ("adaptive mutations") is several orders of magnitude higher than the original RNA cloned from patients.

(B) Modified method

30

Adaptive mutations produced and identified according to (A) can be transferred into a HCV-RNA construct with low ability to replicate. This leads to a huge increase in the replication of this construct. The increase is so great it can be demonstrated that HCV-RNAs transfected into cell culture can replicate even in the absence of selection pressure. Fig. 12 shows a comparison of the replication 35

efficiency of HCV-RNAs, which corresponded either to the starting sequence or to the adaptive sequences 9-13F or 5.1. For the purposes of simple measurement, the neo-gene was removed and replaced by the gene for luciferase. The negative control used was again a HCV-RNA construct that was replication defective due to a deactivating mutation in the NS5B RNA polymerase. Already 24 hours after transfection a noticeable difference is evident in luciferase activity between the defective RNA and the 9-13F or 5.1 constructs, while hardly any difference could be seen between the defective RNA (318 DN) and the parental RNA construct (wt) that possessed no adaptive mutations. During the whole period of observation, the highest luciferase activity, and thus highest replication, was obtained with the 5.1-RNA. These results not only confirm the high replication efficiency of this RNA, but also show that it is possible to create a cell culture system with adapted HCV-RNA constructs for which the presence of a selectable gene is no longer necessary. A summary of the nucleotide and amino acid differences between the starting construct and the mutants 9-13F, 5.1 and 19 is presented in Table 3.

25 **Example 8: Production of cell culture adapted HCV-RNA full-length genome**

In the examples 1 to 7 a subgenomic HCV-RNA was used which lacked the whole structural protein region from core up to 30 p7 or even NS2. It will be shown in this example that it is possible to make a HCV full-length genome replicate in cell culture with the aid of an adapted NS3-5B sequence. For this purpose the *SfiI* fragment of the highly adapted HCV-RNA 5.1 produced according to Example 7 is first transferred into a selectable HCV full-length genome (Fig. 35 12). This HCV genome was transfected into naive Huh-7 cells and subjected to selection with various G418

concentrations. Depending on the strength of selection (the G418 concentration), a varying large number of cell clones was obtained (Fig. 12 B). By contrast no colonies were obtained with the parental HCV full-length genome containing no adaptive mutations, as was the case for the negative control, which was replication defective due to a deactivating mutation in the NS5B RNA polymerase. To confirm that the thus resulting cell clones really contained an autonomously replicating HCV full-length construct, total RNA from several cell clones was isolated and analysed by means of the Northern Blot method. The full-length HCV-RNA was clearly detectable in all cell clones (Fig. 12). It is thus clearly confirmed, that with the aid of cell culture adapted HCV sequences it is possible to produce a HCV full-length genome, which replicates highly efficiently and autonomously in a cell line, i.e. adapted HCV full-length genomes can also be produced with the system of the invention. Furthermore, as this clone possesses the complete HCV sequence, i.e. it also possesses the structural proteins necessary for virus particle formation, it is possible to produce large quantities of infectious virus particles in cell cultures with this system. As a confirmation of these viruses, cell-free cell supernatants carrying a replicating HCV full-length genome, are added to naive Huh-7 cells and the thus infected cells subjected to selection with G418. Each cell clone growing under these conditions originates from an infected cell. The viruses in the cell culture supernatant of cells possessing a replicating HCV full-length genome can be enriched and purified using various known state of the art methods such as ultracentrifugation or microdialysis. They can then be used for the infection of naive cells. Using this method it is clearly demonstrated that cell culture adapted full-length genomes can be produced with the HCV cell culture system of the invention. These genomes replicate with high efficiency in cells and produce infectious viruses. The latter can be detected by

infection of an experimental animal, preferably a chimpanzee.

5 **Example 9: Production of HCV full-length constructs and HCV subgenomic constructs with reporter genes**

A HCV-RNA construct is produced in which a reporter gene is inserted in place of the antibiotic resistance gene (Fig. 13). Replication can thereby be determined through the quantity or activity of the reporter gene or reporter gene product. The reporter gene is preferably a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) or the SEAP gene (secreted alkaline phosphatase gene). This reporter gene and its products, namely the relevant reporter proteins, can be detected for example using fluorescence, chemiluminescence, colorimetrically or by means of immunological methods (for example, enzyme-linked immunosorbent assay, ELISA).

The reporter gene can be expressed either from a separate IRES or in the form of a fusion protein, which is active either as such or fused with a HCV protein via a proteolytically cleavable amino acid sequence in such a way that the reporter is separated from the HCV protein by cleavage of a cellular or viral (HCV) protease.

30

Example 10: Production of HCV full-length constructs with integrated foreign genes used as liver cell specific vectors for gene therapy or as expression vectors

35

The construct (Fig. 14) is transfected in cells and leads to the formation of HCV virus particles that can be used

for the infection of further cells. Since the virus particles have encapsidated RNA with a foreign gene, it can be used in the infected cells for the production of the protein coded by this foreign gene. Cells transfected with the construct also express the foreign gene.

Example 11: Production of monocistronic HCV-RNA constructs in which the resistance gene product is expressed as a fusion protein with the HCV portion.

It is an advantage for some tests if the HCV-RNA construct does not possess a heterologous IRES element. Tests of this type are, for example, the determination of interferon resistance. If a cell possessing a HCV-RNA construct is incubated with interferon alpha or beta, a reduction in replication of the HCV-RNA results. In order to explain the mechanism of this effect it is necessary for the HCV-RNA construct not to possess any heterologous IRES, as otherwise it is not possible to determine whether the interferon mediated inhibition is via inhibition of the HCV replication or inhibition of the heterologous IRES. For this reason constructs are produced for which the resistance gene is fused with a HCV protein (Fig. 15). Either the fusion protein is active as such or the resistance gene product is linked to a HCV protein via a proteolytically cleavable amino acid sequence in such a way that it is separated from the HCV protein by a cellular or viral (HCV) protease.

Table 1: Specific infectivities (cfu/µg RNA) of HCV RNA constructs with adaptive mutations found with the 9-13F mutant introduced into the parental construct I₃₈₉/NS3-3'/wt

amino acid exchange ¹	HCV protein	cfu/µg RNA ²
none		30 - 60
1283 arg -> gly	NS3	200 - 250
1383 glu -> ala	NS3	30 - 60
1577 lys -> arg	NS3	30 - 60
1609 lys -> glu	NS3	160 - 300
(1283 arg -> gly + 1383 glu -> ala + 1577 lys -> arg + 1609 lys -> glu)	NS3	360 - 420
1936 pro -> ser	NS4B	1000-5000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5A	30 - 60
all together		5000

¹ amino acid change in the polyprotein of the HCV isolate con 1 (EMBL-gene bank No. AJ238799); amino acids are given in single letter code.

² Colony forming units (number of cell clones) obtained with a selection of 500 µg/ml G418.

Table 2: Specific infectivities (cfu/ μ g RNA) of the parental HCV RNA construct I₃₈₉/NS3-3'/wt and the variants 9-13I, 9-13F, 5.1 and 19.

Transfected RNA variant	cfu/ μ g RNA ¹
wild type	30 - 50
9-13 I	100 - 1.000
9-13 F	1.000 - 10.000
5.1	50.000 - 100.000
19	50.000 - 100.000

¹ Colony forming units (number of cell clones) obtained with a selection of 500 μ g/ml G418.

Table 3: Nucleotide and amino acid sequence differences between the parental HCV RNA construct I₃₈₉/NS3-3'/wt and the mutants 9-13I, 9-13F, 5.1 und 19.

Clone	nt-position	nt-exchange	aa-exchange
9-13 I	3685	C > T	P > L
	4933	C > T	T > M
	5249	T > C	-
	8486	C > T	-
	8821	G > A	W > stop
	8991	C > G	R > G
	9203	A > G	-
	9313	T > C	F > S
	9346	T > C	V > A
9-13 F	3866	C > T	-
	4188	A > G	R > G
	4489	A > C	E > A
	4562	G > A	-
	4983	T > C	-
	5071	A > G	K > R
	5166	A > G	K > E
	6147	C > T	P > S
	6829	A > G	E > G
	7329	A > G	K > E
NK5.1	7664	A > G	I > V
	8486	C > T	-
	8991	C > G	R > G
	4180	C > T	T > I
	4679	C > T	-
	4682	T > C	-
	5610	C > A	L > I
	6437	A > G	-
	6666	A > G	N > D

	6842	C > T	-
	6926	C > T	-
	6930	T > C	S > P
	7320	C > T	P > S
	7389	A > G	K > E
NK19	3946	A > G	E > G
	4078	C > G	A > G
	4180	C > T	T > I
	4682	T > C	-
	5610	C > A	L > I
	5958	A > T	M > L
	6170	T > A	-
	6596	G > A	-
	6598	C > G	A > G
	6833	C > T	-
	6842	C > T	-
	6930	T > C	S > P
	7141	A > G	E > G
	7320	C > T	P > S
	7389	A > G	K > E
	7735	G > A	S > N

Given are the differences between the nucleotide and amino acid sequences of the parental HCV RNA sequence con 1 (EMBL-gene bank No. AJ238799) and those of the cell culture adapted HCV RNAs. Numbers refer to the nucleotide and amino acid positions of the con 1 isolate. nt, nucleotide; aa, amino acid.

SEQUENCE LISTING

<110> Bartenschlager, Ralf

<120> Hepatitis C Virus cell culture system

<130> ba-1

<140> 199 15 178.4

<141> 1999-04-03

<160> 11

<170> PatentIn Ver. 2.1

<210> 1

<211> 11076

<212> DNA

<213> Hepatitis C Virus

<400> 1

gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 ttttcacgca gaagcgctt agccatggcg ttagttagtg tgcgtgcag cctccaggac 120
 ccccccctcc gggagagcca tagtggctcg cggaaaccgt gagaatcccg gaattgcccag 180
 gacgaccggg tccttcttg gatcaaccgg ctcaatgcct ggagatggg gctgtcccc 240
 gcgagactgc tagcccgatg tggtgggtc gcgaaaggcg ttgtgtact gcctgatagg 300
 gtgcttcgca gtggcccgaa aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
 ctcaaagaaa aaccaaaacgt aacaccaacg ggccgcggcat gattgaacaa gatggatgc 420
 acgcagggttc tccggccgct tgggtggaga ggctattcg ctatgactgg gcacaacaga 480
 caatcggtcg ctctgtatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttctt 540
 ttgtcaagac cgacctgtcc ggtgccttga atgaactgtca ggacgaggca ggcggcgat 600
 cgtggcttgc cacgacggc gttcccttgcg cagctgtgtc cgacgttgc actgaagcg 660
 gaagggactg gctgttattc ggcaagtgc cggggcaggag tctctgtca tctcacctt 720
 ctcttcggca gaaatgttcc atcatggctg atgcataatgcg gccgtgtcat acgcttgatc 780
 cggctactcg cccatcgac cacaaggca aacatcgatcg cgtactcgaa 840
 tggaaaggccg tcttgcgtat caaggatgtc tggacgaaaga gcatcagggg ctcgcgcag 900
 ccgaactgtt cggccaggctc aaggcgcgc tggccgcgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgtttgcg aatatcatgg tggaaaatgg ccgttttctt ggattcatcg 1020
 actgtggccg gctgggtgtc ggggaccgt atcaggacat agcgttggc acccgatata 1080
 ttgtcaaga gcttggccgc gaatgggtcg accgttctt cgtgttttac ggtatcgccg 1140
 ctcccgatc gcacggcattc gcttttcttgc gcttcttgcg ctagtttctc tgagttaaa 1200
 cagaccacaa cggtttccctt ctggccggat caatttcggc cctcttcctc cccccccct 1260
 aacgttactg gccgaaaggcg ctggaaataa ggccgggtgtc cttttgtcta tatgttattt 1320
 tccaccatat tgccgttctt tggcaatgtg aggccggaa aacccggcc tttttttttt 1380
 acgacgttcc ctgggggtct ttcccctctc gccaaaggaa tgcacggatc gttgaatgtc 1440
 gtgaaggaaag cagttccctt ggaagcttct tgaagacaaa caacgttgc gtcgttgc 1500
 tgcaggcgc ggaaccccccc acctggcgac aggtgcctct gggccaaaa gccacgtgt 1560
 taagatacac ctgcaaaaggc ggcacaaccc cagtgccacg ttgtgagtt gatagttgt 1620
 gaaagagatca aatggctctc ctcaagcgta ttcaacaagg gggtgaagga tgcccgaaag 1680
 gtatccccattt gtatggatc tgatctgggg cctcggtgc catgttttac atgtgtttat 1740
 tcgagggttaa aaacatgtca ggccccccca accacggggc cttttttttt ctttggaaaa 1800
 cacgataataa ccatggccac gaatccaaatgg cttccaaagaa aacccaaacg taacaccaac 1860
 cggccggccac aggacgtcaa gttccggcggc ggtggcgtcgc tcgtcggtgg agtttacctg 1920
 ttggccgcgc gggggccca gttgggtgtc cgccgcacta ggaagacttc cgagcggtcg 1980
 caacctcgtaa gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggccttg 2040
 gtcagcccc ggttcccttgc gcccctctat ggcaatgggg gcttgggtgtc ggcaggatgg 2100
 ctctctgtcact cccgtggctc tggccctagt tggggccca cggccccccg gctgtggatc 2160
 cgcacattttt gtaaggatcat cgtatccctc acgtgcggct tcggcgatct cttttttttt 2220
 attcccgctcg tcggccccc ctttggggc gtcggccaggc ccctggcgca tggcgccgg 2280
 gtttggggagg acggcgatc ctatgcataa gggaaatgcg cccgttgc cttttttttt 2340
 ttcccttttttgc ctggatgtc ctgttttgacc atcccaatgg cccgttatgt agtgcgcac 2400
 gtatccggag ttttccatgtt caccggatcg tgcgtccaaatgg caagcattgtt gttatggac 2460
 ggggacatga tcatgcatac ccccggtgc gtcggccatgg tttttttttt cttttttttt 2520
 cgtctgtgg tagcgatc tccacgttc gggccaggaa acgttagctt ccccaactacg 2580
 acgatatacgac gccatgtcgat ttttccatgtt cccgttgc cttttttttt cttttttttt 2640
 gtggggatc ttttccatgtt cccgttgc cttttttttt cttttttttt cttttttttt 2700
 cggccacgaga cttttttttt cttttttttt cttttttttt cttttttttt cttttttttt 2760

cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgtag 2820
 ttactccgga tccccacaagc tgcgtggat atggggcgg gggcccatgg gggagtccta 2880
 gcgggccttg ctactattc catggtggg aactggcta aggttctgat tgtatgcta 2940
 ctcttgcgg gcgttgcgg gggacccatgt gtgacagggg ggacatggc caaaaacacc 3000
 ctcgggatta cgtccctttt ttcacccggg tcatcccaga aaatccaaatgt tgtaaacacc 3060
 aacggcagct ggacatcaa caggactgca ctgaactgca atgactccct caacactggg 3120
 ttcccttgctg cgctgttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcata 3180
 gcccacgtca gccccatcgca cgcgttcgtc caggggtggg gggccatcac ttacaatgag 3240
 tcacacagct cgAACAGAG gccttattgt tggcactacg caccggggcc gtgcgttac 3300
 gtaccggcg cgcaagggtgt tggtccatgt tactgttca ccccaagccc tgctgttgt 3360
 gggacgaccg accgggttcgg cggtccatcg tacagggtgg gggagaatgt gacggacgtg 3420
 ctgcttcta acaacacgcg gcccggcaaa ggcaactgtt ttggctgtac atggatgaaat 3480
 agcaactgggt tccaacaaatgt gtcgggggc cccctggta acatgggggg gatcgcaat 3540
 aaacccctgaaatgtt cctggccccc cggatgttcc cggaaacacc cccggggccat ttacaccaag 3600
 tggtgttgcg ggccttggtt gacacccaga tgcttgcgtt actacccatata caggcttgg 3660
 cactacccct gcaactgtcaa cttaaccatc ttcaaggta ggtatgtacgt ggggggagtg 3720
 gagcacaggc tcgaagccgc atgcaatttg actcgaggag agcgttgtaa cctggaggac 3780
 agggacagat cagacttag cccgtctgt ctgtctacaa cggagtgcc ggtattgccc 3840
 tgttccttca ccacccatcc ggctctgtcc actgggttga tccatcttca tcagaacgtc 3900
 gtggacgtac aatactgtt cggatataagg tcggcggttgc tctcccttgc aatcaaatgg 3960
 gagatgttcc tggtgttctt ccttcttgc gggacggcgc gctgtctgtc ctgtttgtgg 4020
 atgatgttgc tgataatgtca agctggggcc gcccataagc acctgggtgtt cctcaacgcg 4080
 gcatccgtgg cccggggcga tggcatttttcc tcccttcgt tggttttgc tgctgcctgg 4140
 tacatcaagg gcaggttgcggtt ccctggggccg gcatatgccc tctacggcgt atggccgtt 4200
 ctccctgttcc tgctggcggtt accaccacgca gcatacgcca tggacggggaa gatggcagca 4260
 tcgtgcggag gcgcgggtttt cgttaggtctg atactcttgc cttgttaccg gcaactataag 4320
 ctgttccctcg cttagtctcat atgggtgttca caatatttta tcacccaggc gggggccat 4380
 ttgcgaatgtt ggttttttttcc cctcaacgtt cggggggggcc gcatgtccgtt catccttc 4440
 acgtgcgcga tccacccaga gctaatttttcc accatcacca aaafcttgcg cggccataactc 4500
 ggttccactca tggtgttccca ggttggata accaaatgtc tgtagtgcgt ggcgcacac 4560
 gggcttattc tgcatgtcat gttgtgcgg aagtttgcgtt ggggttatttgc tgccaaatgt 4620
 gctctcatgtt ggttttttttcc cctcaacgtt cttgttgcgtt cagttgagcc ggttgcctgg 4680
 cggggactggg cccacgcggg cttacggac gttgttgcgtt cttgttgcgtt cgttgccttc 4740
 tctgtatgg agaccaaggat tatcacctgg gggggcagaca cggcggcgtt gttggacatc 4800
 atcttggcc tggccgttcc cggccgcagg gggagggaga tacatctgg accggcagac 4860
 agcccttgcgag ggcgggggtt ggcacttcc ggcgttatttgc cggcttactc ccaacagac 4920
 cggggcttac ttggctgtcat catcaatgtt ctcacaggcc gggacaggaa ccaaggctc 4980
 gggggagggttcc aatgggttcc caccgcacca caatatttcc tggcgttccgt cgtcaatggc 5040
 tggtgttgttgc tggcttatca ttggcggccggc taaaagcccc ttggccggccaa aaggggccca 5100
 atcacccaaa tgacacccaa tggtggaccat gacctgttgc gttggcgttccgttcc 5160
 ggcgttcttgc tgacaccatg cacctgcggc agctcgacc ttacttgcgttccgttcc 5220
 ggcgttcttgc tttccgttgcg cccggggggcc gacagcaggg ggagcctacttcc 5280
 cccgttcttgc acttgcagggttcc tttccgttgcg ggttccacttc tctcccttc 5340
 gtggggcatct ttccggctgc cgtgtgcacc cgggggttgc cgaaggccgtt ggttttgc 5400
 cccgttgcgttcc tttccgttgcg cttccgttgc tttccgttgc tttccgttgc 5460
 ctatggaaac cactatgttgc tttccgttgc tttccgttgc tttccgttgc 5520
 cccggccgttcc cccggccgttcc cccggccgttcc cccggccgttcc 5580
 cccggccgttcc cccggccgttcc cccggccgttcc cccggccgttcc 5640
 cccggccgttcc cccggccgttcc cccggccgttcc cccggccgttcc 5700
 aacatcagaa cccgggttgc gacccatcacc acggggccccc cccatcacgttcc 5760
 ggcaagtttgc ttggccacccggg tttccgttgc tttccgttgc tttccgttgc 5820
 gatgttgcacttcc caactgttgc gacccatcacc acggggccccc cccatcacgttcc 5880
 gagacggccgttcc gacccatcacc acggggccccc cccatcacgttcc 5940
 gtggccacccatc cccatcacgttcc cccatcacgttcc cccatcacgttcc 6000
 ggcggccgttcc cccatcacgttcc cccatcacgttcc cccatcacgttcc 6060
 tattaccggg gcttgcgttcc tttccgttgc tttccgttgc tttccgttgc 6120
 acggacgttcc tttccgttgc tttccgttgc tttccgttgc tttccgttgc 6180
 tggtgttgc tttccgttgc tttccgttgc tttccgttgc tttccgttgc 6240
 gtcggccatccatc acggccgttcc tttccgttgc tttccgttgc tttccgttgc 6300
 ggcatttttccatc tttccgttgc tttccgttgc tttccgttgc tttccgttgc 6360
 ctgttgcgttcc tttccgttgc tttccgttgc tttccgttgc tttccgttgc 6420
 gtttaggttgc tttccgttgc tttccgttgc tttccgttgc tttccgttgc 6480
 ttcttggggatc gcttgcgttcc tttccgttgc tttccgttgc tttccgttgc 6540
 aaggccggccatc gacccatcacc acggccgttcc tttccgttgc tttccgttgc 6600
 gcttgcgttcc tttccgttgc tttccgttgc tttccgttgc tttccgttgc 6660
 acggacgttcc tttccgttgc tttccgttgc tttccgttgc tttccgttgc 6720
 accacacaccatc cccatcacgttcc tttccgttgc tttccgttgc tttccgttgc 6780

acggcacct ggtgctgtt aggccggatc ctagcagctc tggccgcgt a 6840
acaggcagcg tggcattgt gggcaggatc atcttgtccg gaaagccgc cátcatccc 6900
gacagggaaag tccttaccg ggagttcgat gagatggaaag a 6960
tacatcgaaac a 6960
caaacaacgcca ccaagcaagc ggaggctgt gtccttgcgtt gaaatccaa gttgcggacc 7080
ctcgaagect tctggcgaa gcatatgtgg aatttcatca gcccggataca atatttagca 7140
ggcttgcctca ctgccttgcg caacccccggc atagcatcac tgatggccat cacagctct 7200
atcaccagcc cgctcaccac ccaacatacc ctcccttta acatcttgcg gggatgggt 7260
gcccggccaaac ttgtcttcc gäggctgtgg gtccttgcgtt a 7320
gccccgttgg gcaagcatagg cttggggatgg a 7380
gccccggcgtt gggacactgg ttaacctact cttggggatgg a 7440
gtgtgcgcag cgatactgcg ccctgtatc ctctccctg gcgccttgcgtt 7500
a 7560
a 7620
a 7680
a 7740
a 7800
a 7860
a 7920
a 7980
a 8040
a 8100
a 8160
a 8220
a 8280
a 8340
a 8400
a 8460
a 8520
a 8580
a 8640
a 8700
a 8760
a 8820
a 8880
a 8940
a 9000
a 9060
a 9120
a 9180
a 9240
a 9300
a 9360
a 9420
a 9480
a 9540
a 9600
a 9660
a 9720
a 9780
a 9840
a 9900
a 9960
a 10020
a 10080
a 10140
a 10200
a 10260
a 10320
a 10380
a 10440
a 10500
a 10560
a 10620
a 10680
a 10740
a 10800

ctactttctg taggggtagg catctatcta ctccccaacc gatgaacggg gagctaaaca 10860
ctccaggcca ataggccatc ctgtttttt ccctttttt ttttctttt tttttttttt 10920
ttttttttt ttttttttc tcctttttt ttccctttt ttccctttc ttccctttgg 10980
tggtccatc ttggccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040
agagagtgtct gatactggcc tctctgcaga tcaagt 11076

<210> 2
<211> 8637
<212> DNA
<213> Hepatitis C Virus

<400> 2

```

gcccggcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tettcacgc aaaaacgtct agccatggcg ttatgtatgag tgcgtgcag cctccaggac 120
ccccccctcc gggaaagcca tagtggctcg cggaaaccgt gagaatccacccg gaattgcccag 180
gacgaccggg tccttcctt gatcaaccgg ctcataatgcct ggagatggg gcgtgcccc 240
gcaaaaaactgc tagccggat gtgtggggc gcaaaaggcc ttgtggtaact gctgtatagg 300
gtgcttgcgat gtgcggccgg agggttcgtat gacccgtcact catgagcacg aatcctaaac 360
ctcaaaagaaa aaccaaaaggg cgccggcatgtt tgaaacaaga tggatgcac gcagggttctc 420
cggccgcttg ggtggagaggg ctatccggct atgactggc acaacacgaca atcggctgtct 480
ctgatgcgcg cgtttccgg tcgtcagcgc agggccccc ggttctttt gtcaagacgg 540
acctgtccgg tgcctgtat gaaactgcagg acgaggcgc goggctatcg tggctggcca 600
cgacggggct tccttgcgca gctgtgtcg acgttgtcac tgaagcgggaa agggacttgc 660
tgcttattggg cgaagtgcgg gggcaggatc tcctgtatcc tcaccttgc tggccgaga 720
aagtatccat catgtgtat gcaatgcggc ggctgcataatc gctgtatccg gctacccgtc 780
catccgacca ccaacggaaa ctcgcacatc acggacacg tactccggatc gaaagccgtc 840
ttgtcgtatca ggtatgtatcg gacgaagagc atcagggtct cggccggatcc gaaactgttcg 900
ccaggctcaa ggcgcgcatcg cccgacggcg aggatctcg ctgtgacccat ggcgtatcc 960
gcttgcggaa tatcatgggt gaaaatggcc gcttttctgg attcatcgatc tggccggc 1020
tgggtgtggc ggaccgttat caggacatacg cgttggatc cgtgtatatt gctgaagagc 1080
ttggcggcgat atggctgac cgttctctcg tgcttacgg tattgcgcgt cccgatttgc 1140
agcgcatcgc ctctatcgc ttcttgcacg agttcttctg agtttaaaca gaccacaacg 1200
gttccctctt agccggatcatc ttccggccccc ttcctccccc ccccccctaa cgttacttgc 1260
cgaaggccct tggaaataagg cgggtgtgcg ttgttataatc tggccgttgc caccatattt 1320
ccgtcttttgc gcaatgttag gggccggaaa cctggccctg tcttgcattt gaggatttgc 1380
aggggtcttt cccctctcgc caaaggaaatg caaggctctg tgaatgtcg gaaaggacg 1440
gttccctctgg aagttcttg aagacaaaca acgtctgtatc cggccctttt caggcagcgg 1500
aacccccccac ctggcgacag gtgcctctgc ggcggaaatc cacgtgtata agataacacct 1560
gcaaggccgg cacaacccca gtgccacgtt gtgagttggatgttgcg aagagtcaaa 1620
tggctctctt caagcgattt caacaagggg ctgaaggatg cccagaaggat gacccatttt 1680
atgggatctg atctggggcc tcgggtgcaca tgcttacatc gtgttttagtc gaggttaaaa 1740
aacgtctagg ccccccgaac cacggggacg tggtttctt ttgaaaaca cgtataatacc 1800
atggacccgg agatggcgc atcgtgcggg ggcggccgtt ctgtggatc gataacttgc 1860
accttgcac cgcactataa gctgttccctc gctaggctca tattgtgtttt acaatatttt 1920
atcaccaggc ccgaggcaca ttgcacgtt tgatcccccc cctcaacatc tcgggggggc 1980
cgcgatggcc tcatctctt cactgtgcgc atccacccag agctaatttt taccatcacc 2040
aaaatcttgc tcgcataact cggtccactc atgggtctcc aggtgttat aaccaaaatgt 2100
ccgtacttgc tgccgcaca cgggtcttatt cgtgcacatc tgctggatc gaaagggttgc 2160
gggggttattt atgtccaaat ggctctcatg aagttggccg cactgacagg tacgtacgtt 2220
tatgaccatc tcacccactt cggggacttgc gcccacggc gcctacgaga ctttgcgggt 2280
gcagtttgacg cctgtctttt ctgtatatgc gggccaaatc ttatcacctt gggggcagac 2340
accggccggat ctttggggatcatc ttggccgttcc cccggccggc gggggagggg 2400
atacatctgg gacccgcaga cagcccttgc gggcagggtt ggcacttctt cgcgcctatt 2460
acggcctact cccacacgac ggcggccata cttggctgc tcatcactatg cctcacaggc 2520
cgggacacgg accaggtcg gggggagggtc caagtgttcc ccacccgac acaatcttt 2580
ctggcgaccc tgcgtcatgg cgtgtgttgc actgtctatc atgtgtccgg ctcaaagacc 2640
cttgcggccca cccaaaggccca aatcacccaa atgtacacca atgtggacca ggacccgtc 2700
ggctggcaag cggcccccgg ggcgggttcc ttgcacccat gcacccgtgg cagctggac 2760
ctttaacttgc tcacggggca tgccgtatgc atccgggtc gcccgggggg cgacacggcagg 2820
ggggacccatc tctcccccgg cggccgttcc tactgtatgc gcttccgggg cggtccactg 2880
ctctggccccc cggggcggcc tttggggatc tttggggatc cccgtgtgcac cccgggggtt 2940
gcgaaggccgg tggactttgtt accccgtcgat tctatggaaa ccactatgc gttccgggtc 3000
ttcacggaca actcgcccccc tcggccgtatcc cccggatccat tccagggtgc ccatctacac 3060
gcccctactg gtatggccaa gggccatcaag gtgcgggtc cgtatgcac ccaagggttat 3120
aagggtcttgc tcctgtatcc gtcgtgcgc gcccacccatc gtttggggcc gatatgtct 3180
aaggccatcg gtatcgaccc taacatcaga accggggtaa ggacccatc cccgggtgcc 3240
cccatcactt actccacca tggcaagttt cttggccgtc gttgtgttc tggggggcc 3300
tatgacatca taatatgtatgc tgatgttgcac tcaactgtact cggccatcatc cttggggccatc 3360
ggccacatcg tggacccatc gggacccggc gggccggccatc tcgtgtgc gcccacccgt 3420
acgcctccgg gatcgatcgc cgtggccatcatc cccaaatcg agggaggtgc tctgtccacg 3480
actggagaaa tcccccttta tggcaagcc atccccatcg agacccatcaa gggggggagg 3540
cacctcattt tctgccccatc caagaagaaa tggatgtatgc tggccggc 3600

```

ctcggactca atgtgttagc atattaccgg ggcctgatg tatccgtcat accaactagc 3660
 ggagacgtca ttgtcgtagc aacggacgtc ctaatgacgg gcttaccgg cgatttcgac 3720
 tcagtgtacg actgcaatac atgtgtcacc cagacagtgc acttcagccct ggacccgacc 3780
 ttcaccattg agacgacgac cgtgccacaa gacgcgggtgt cäcgctcgca gcggcgaggc 3840
 aggactgtta gggcaggat gggcatttac aggttgtga ctccaggaga acggccctcg 3900
 ggcatgtcg attctcggt tctgtcgag tgctatgacg cggcgtgtgc ttggtagcag 3960
 ctcacgccc ccgagaccc ägttaggtt cggcgttacc taacacacc ägggttcccc 4020
 gtctgcäagg accatctgga gttctgggag agcgttta caggcctcac ccacatac 4080
 gcccattct tgcctccagac taagcaggca ggagacaact tcccctaccc ggtacatc 4140
 caggctacgg tgcgcgcag ggctcaggct ccacccat cgtgggacca aatgtggaa 4200
 tgcctcatac ggctaaagcc tacgtgcac gggcaacgc ccctgtgtta taggctggg 4260
 gccgttcaaä acgagggtac taccacacac ccataäcca aatacatcät ggcacatgc 4320
 tcggctgaccc tggaggctgtt caccggcacc tgggtgtgg taggcggagt cctägcac 4380
 ctggcccggtt attccgtac aacaggcgcgt gtgttcattt gggcggatcatcttgc 4440
 ggaaaggccgg ccatccatcc cgcacgggaa gtccttacc gggagttcga tgagatgaa 4500
 gagtgccctt cacacccctt tacatcgaa caggaaatgc agctcgccgaa acaattcaa 4560
 cagaäggcaa tcgggttgc gcaaacagcc accaagcaag cggaggctgc tgctccctg 4620
 gtggaatcca agtggcgac ctcgaagcc ttctggcga agcatatgtg gaatttcatc 4680
 agcgggatac aatatttagc aggttgtcc äctctgcctg gcaäccccgc gatagcatca 4740
 ctgatggcat tcacagcctc tattaccaggc ccgtctacca cccaacatac cctccctgtt 4800
 aacatcttggg gggatgggt ggccgccccaa cttgtctcc ccaacgcgtc ttctgtttc 4860
 ttaggcggcc gcatcgctgg agcgggtgtt ggcacatag gcttgggaa ggtgttgtg 4920
 gatattttgg cagggtatgg acgggggtt gcaaggcgcgc tctgtggcctt taaggctatg 4980
 agcggcgaga tgccctccac cgaggacttgc gttaaatcttac tccctgttat cctccctt 5040
 ggcgccttag tcgtcggtt cgtgtgcga gcgataactgc gtcggcacgt gggcccaagg 5100
 gagggggctg tgcagtggat gaaacggctg atagcgttc cttcgccggg taaccacgc 5160
 tccccccacgc actatgtgcc tggagacgcac gtcgcacac gtgtcactca gatcctct 5220
 agtcttacca tcaactcagct gctgaagagg cttcaccatg ggtacaaacgaa gactgcctc 5280
 acgcatatgc ccggctcggt gctaagaagat gtttggatt ggätatgcac ggtgttgc 5340
 gatttcaaga cctgctcca gtcacagctc ctgcggcgt gtcggggagt ccccttctt 5400
 tcatgtcaac gtgggtacaa gggagtctgg cggggcgcac gcatcatgca aaccacctgc 5460
 ccatgtggag cäcagatcäc cgcacatgtg aaaaacgggtt ccatgaggat cgtggggctt 5520
 aggacctgtt gtaacaägtg gcatggaaca ttccccatca acgcgtacac cacggcccc 5580
 tgcacgcctt cccggcgcc aaattattctt agggcgtgtt ggcgggtggc tgctgaggag 5640
 tacgtggagg ttacgcggg ggggatttc cactacgtgä cgggcägtgac cactgacac 5700
 gtaaaagtgcc cgtgtcagg tccggccccca gaatttctca cägaagtggaa tgggtgcgg 5760
 ttgcacaggt acgctccagc gtgcaaaacc ctcctacggg aggagggtcac attcctgtc 5820
 gggctcaatc aatacttggg tgggtcacag cttccatgcg agcccgaaacc ggacgtagca 5880
 gtgctcaattt ccacatgcac cgcacccctcc cäcattacgg cggagacggc taagcgtägg 5940
 ctggccagg gätcctccccc ctccctggcc agtcatcag ctacgcac gtcgtccct 6000
 tccttgaagg caacatgcac taccgtcat gactccccc acgctgaccc catcgagcc 6060
 aacctcctgtt ggcggcagga gatggggggg aacatcaccac ggtggagtc agaaaataäg 6120
 gtagtaattt tggactctt ctagccgctc caäcgggagg aggtatgagag ggaagtatcc 6180
 ttcccgccgg agatcctgcg gagggtccagg aaattccctc gagcgatgcc catatggca 6240
 cggccggatt acaaccctcc äctgttagat tccctggaaagg accggacta cgtccctcca 6300
 gtggtacacg ggtgtccat ggcgcctgc aaggccctc gcatäaccacc tccacggagg 6360
 aagäggacgg ttgtctgtc agaatctacc gtgttcttc ctttggggaa gtcgcacaca 6420
 aagacccctcg gcaactccgä atcgtcgcc gtcgcacgc gcaacggcää ggcctctt 6480
 gaccäccctcg cäcagacgg cgcacgggaa tccacgttg agtcgtactc ctccatgccc 6540
 ccccttggagg gggagccggg ggtatcccgtt ctcacgcacg ggttggcgtt taccgtaaac 6600
 gaggaggctt gtagggäcgä cgtctgtgc tcgatgttcc acacatggac aggccctcg 6660
 atcacgcacat ggcgtcgacca gggaaaccaac ctggccatca atgcactgag caactctt 6720
 ctccgttacc acaacttggg ctatgttacca acatctcgac ggcacggccct gcccggaga 6780
 aagggttaccat ttgacacccat gacgttgcctg gacgaccact accgggacgt gctcaaggag 6840
 ätgaaggöga aggcgtccac agttaaggat aaacttctat cctgtggaggaa agcctgtaaag 6900
 ctgacgcccc cacatcgcc cagatctaaa ttggatgtt gggcaaaaggaa cgtccggaaac 6960
 ctatccagca aggccgttac ccacatccgc tccgtgtggä aggactgtt ggaagacact 7020
 gagacaccat ttgacaccat catcatggca aaaaatgagg ttttgcgtt ggcacggcc 7080
 aaggggggcc gcaaggccac tccgccttacat gtttccatc atttgggggt tcgtgtgtc 7140
 gagaaaaatgg cccttacga tgggtctcc accctccctc äggccgttat gggctcttca 7200
 tacggattcc aatactctcc tggacacgcgg gtcgaggatcc tggtagatgc ctggaaagcg 7260
 aagaaatgcc ctatggctt cgcacatgcac accccgtgtt ttgactcaac ggtcactgag 7320
 aatgacatcc gtttgcgtt gtcacatctac caatgtgtg acttgcccc cgaaggccaga 7380
 cäggccatcaa ggtcgctcac agagcggctt tacatcggg gcccctgc taattctaaa 7440
 gggcagaact ggcgttactcg ccgtgcgcgc gcgacgggtg tactgacgac cagctgcgg 7500
 aataccctca catgttactt gaaggccgtt gccgcctgtc gaggtgcgaa gtcgcaggac 7560
 tgcacgtacg tgcgtatgcgg agacgaccctt gtcgttatct gttggaaacccaa 7620

ggggacgagg	cgaggctacg	ggccttcacg	gaggctatga	ctagatactc	tgcggccctt	7680
ggggaccgc	ccaaaccaga	atacgacttg	gagttgataa	catcatgtc	ctccaatgtg	7740
tcaagtgcgc	acgatgcac	tggcaaaagg	gtgtactatc	tcacccgtga	ccccaccacc	7800
cccccttgcgc	gggctgcgtg	ggagacagct	agacacactc	cagtcaattc	ctggctagc	7860
aacatcatca	tgtatgcgc	cacccctgtgg	gcaaggatga	tcctgtatgc	tcattttctc	7920
tccatccctc	tagctcaggaa	aaacacttgc	aaaggccctag	attgtcatg	tcacggggcc	7980
tgttacttcc	tttagccact	tgacctacct	cagatcattc	aacgactcca	tggccattac	8040
gcattttcac	tccatagttt	ctctccagggt	gagatcaata	gggtggcttc	atgcctcagg	8100
aaaccttgggg	taccggccctt	gcgagttctgg	agacatcggg	ccagaagtgt	ccgcgttagg	8160
ctactgtccc	aggggggggag	ggctgccact	tgtggcaagt	acctttcaa	ctggcgagta	8220
aggaccaagc	tcaaactcac	tccaatcccg	gctgcgtccc	agtggattt	atccaaatgtgg	8280
ttctgttgcgt	tttacagcgg	gggagacata	tatcacagcc	tgttctgtgc	ccgaccccccgc	8340
tggttcatgt	gggtgcctat	cctactttct	gtagggttag	gcatctatct	actcccaaac	8400
cgtgaacgg	ggagetaaac	atccaggcc	aataggccat	cctgtttttt	tccctttttt	8460
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tccttttttt	tttcctcttt	8520
tttttttttt	ctttcccttt	gtggctccat	cttagcccta	gtcacggcta	gctgtgaaag	8580
gtccgtgagc	cgcttgactg	cagagagtgc	tgatacttgc	ctctctgcag	atcaagt	8637

<210> 3
<211> 8001
<212> DNA
<213> Hepatitis C Virus

<400> 3

gccagcccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tettacacgcga aaagcgtct agccatggcg ttatgtatgag tgcgtgcag cctccaggac 120
ccccccctccca ggagagccaa tagtggctcg cggAACCGGT gaatCACCCG gaattgcac 180
gacgaccggg tccttccttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcatggatgc tagccgagta gtgttgggtc gcgaaggcc ttgtggtaact gcctgatagg 300
gtgttgcga gtggcccccggg aaggctcgta gaccgtgcac catgagcacg aatctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggccgcgcatt gattgaacaa gatggatgc 420
acgcagggttc tccggccgt tgggtggaga ggctattcgat ctatgactgg gcacaacaga 480
caatcggtcg ctctgtatgcc gccgtgttcc ggctgtcgcg qcaggggcgc ccggttctt 540
ttgtcaagac cgacctgtcc ggtccctga atgaactgca ggacgaggca ggcggcgtat 600
cgtggctggc cacgacgggc gttccctgca cagctgtct gcacgttgc actgaagcgg 660
gaagggactg gctgttattt ggcgaagtgc cggggcagga tctccctgtca tctcacctt 720
ctctccgcga gaaatgtatcc atcatggctg atgcaatgcg gccgtgtcat acgcttgatc 780
cggacttgc cccatcgac cacaaggca aacatcgat cggacggcga cgtactcgga 840
tggaaagccgg tcttgcgtat cggatgtatc tgacgacaa gcatcagggg ctcgcgcag 900
ccgaactgtt cggcaggctc aaggcgcgcg tggccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgttgcgg aatatacatgg tggaaatgg ccgttttctt ggattcatcg 1020
actgtggccg gctgggtgtg ggggaccgcg atcaggacat acgcttggtt acccggtata 1080
ttgtcaaga gcttggccgc gaatgggctg accgcttcct cgtgtttac ggtatcgccg 1140
ctcccgattc gcagcgcatac gccttctatc gccttcttgc caggttctc tgagttaaa 1200
cagaccacaa cgggttccctt ctggccggat caattccgc cctctccctc cccccccct 1260
aacgttactg gccgaagccg ctggaaataa gggccgtgt cgtttgtcta tatgttattt 1320
tccaccatata tggcccttgc tggcaatgtg aggggccgaa aacccggcc tggcttctt 1380
acgagcattc ctgggggtct ttcccctctc gccaaggaa tgcaaggctt gttaatgtc 1440
gtgaagggaa cagttctctt ggaagcttct tgaagacaaa caacgtctgt agcggccctt 1500
tgcaggcagc ggaacccccc acctggcgac aggtgcctt gggccaaaa gccacgtgta 1560
taagatacac ctgcaaaaggc ggcacaaccc cagtggccacg ttgtgagtt gatagttgt 1620
gaaagactca aatgtctctc ctcaagcgta ttcaacaagg ggctgaaagg tgcccagaag 1680
gtacccctt gtatggatc tgatctgggg cctcggtgca catgtttac atgtgttttag 1740
tgcagggtta aaaacgtcta gcccccccgaa accacggggg cgtgggtttc ctttggaaaa 1800
cagcataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttgc 1860
tgcatacatca ctggctcac aggccggggc agggacccagg tgggggggg ggtccaaatgtc 1920
gtctccaccc caacacaatc tttctggcg acctcggtca atggcgtgt tgactgtc 1980
tatcatgggt cgggtcaaa gacccttgcc gggccaaagg gcccataatc ccaaataatgtc 2040
accaatgtgg accaggaccc cgtcggtcttgg caagcggccc cggggcgccg ttcccttgaca 2100
ccatgcaccc gggcagctc ggacctttac ttgttgcacga ggcataccgca tgcattccg 2160
tgcggccggc gggggccacag caggggggcact ctaactctcc ccaggcccgt tgccaaatgtc 2220
aagggtctt cgggggttcc actgtctgc ccttcggggc acgtgtgttcc ctagtttccg 2280
gtcgccgtgt gcacccgggg ggttgcgaag cgggtgtact ttgatccctt ctaggtgc 2340
gaaaccacta tgcggcccccc ggttcttgc gacaactctt ccctccggc taaagggtccg 2400
acattccagg tggcccatctt acacgccccctt actggtagcg gcaagagccg cggccacc 2460
gctgcgtatg cagcccaagg gtataagggtt ctgttgcacc acccggtccgt tgactgtc 2520
ctaggtttcg gggcgatata gtcataaggca catggatctg accctaatacat cagaaccggg 2580
gtaaaggacca tcaccacggg tggcccatcc acgtactcca cctatggca gtttcttgc 2640
gacgggtggg gctctggggg cgccatgtac atcataataat gtgtatgtt ccactcaact 2700
gactcgacca ctatcttggg catcgccaca gtccctggacc aagccggagac ggctggagcc 2760
cgactcgctg tgctcgccac cgtactcgctt ccgggtatcc tcacccgttcc acatccaaac 2820
atcgaggagg tggctgttc cggactgtt gaaatccctt tttatggca aggccatcccc 2880
atcgagacca tcaagggggg gaggcacctt atttctcc atttcaagaa gaaatgtgt 2940
gagctcgccg cgaagctgtc cggccctcgga ctcaatgtt tagcatatta cggggccctt 3000
gtatgtatccg tcataccaac tagcggtggac gtcattgtcg tagacaacggc cgctctaatt 3060
acgggcttta cggcgattt cgaactcgatg atcgactgca atacatgtt caccctggaca 3120
gtcgacttca gctgtggacc gaccccttccat attgagacga cggccgtcc acaagacgcg 3180
gtgtacgtt cgcagccggc aggccggact ggtaggggca ggtatggcat ttacagggtt 3240
tgactccag gaaacggcc ctggggcatg ttccatgttcc cgggttctgt cgactgtctat 3300
gacgcgggctt gtgttggta cggactcgat cccggcaga cctcgttag gttcggggtt 3360
tacctaaaca caccagggtt gcccgtctgc caggaccatc tggatgttcc gggagggcgtc 3420
tttacaggcc tcaccacat agacgccccat ttcttgcctt agactaagca ggcaggagac 3480
aacttccctt acctgtatc ataccaggct acgggtgtcg ccagggtctca ggctccac 3540

ccatcggtgg accaaatgtg gaagtgtctc atacggctaa agcctacgc
acggccccgc tttataggct gggagccgtt caaaacgagg ttactaccac
accaaataca tcatggcatg catgtcggtt gacctggagg tcgtcacag
ctggtaggcg gactctggcc gctattgcc tgacaacagg
atgtggca ggatcatctt gtccggaaag ccggccatca ttcccacag
taccgggaatgt tcgtcgatg gaaagagtgc gcctcacacc tcccttatcat
atgcagctcg ccgaacaatt caaacagaag gcaatcggt tgcgtcaaac
caaggcgagg ctgtctcc cgtgggtggaa tccaatgtgg gggccctcg
gcaaggatata tggatattt catcaacggg atacaatatt tagcaggctt
ctggcaacc ccggatagc atcaactgtt gcttacacag cctctatcac
accacccaaac ataccctctt gtttacatc ctgggggat gggggccgc
ctcccaagcg ctgtctcg ttcgttaggc gccggcatcg ctggagccgc
ataggcttg ggaagggtgt tggatattt ttggcagggtt atggagcaagg
gcgtcggtt ccttaagggt catgagccgc gagatgcctt ccacccgagga
ctactccctg ctatctctc ccctggccgc ctgtcgctg gggctgtgt
ctcgctggc acgtggggcc aggaggaggg gctgtcgatg ggatgaaccg
ttcgcttcg gggtaaccat ctgtcccccc acgactatg tgcgttagag
gacgtgtca tcagatctt ctctgttaccatc acgtgtcaa
cagttggatca acgaggactg ctccacgcgc tgctcccgct cgtggctaaag
gattggatat gcacgggtt gactgattt aagacctggc tccagttccaa
cgattgcgg ggttccctt ttctctatgt caacgtgggt acaaggaggt
gacggcatca tgcaaccac ctggccatgt ggagcacagaa tcacccgac
ggttccatga ggatcggtt ggcttaggacc ttagtaaca cgtggcatgg
attaacgcgtt acaccacggg cccctgcaccc ccctcccg cgccaaatta
ctgtggccgg tggctgtga ggagtacgtt gagttacgc ggggtgggg
gtgacggggca tgacactgtt caacgtaaag tgccctgtc aggttccggc
ttcacaagaaatggatggggg ggggtgtc acggtacgtc cagcgtgc
cgggaggagg tcacattctt gtcgggctc aatcaatacc tgggggg
tgcgagcccg aaccggacgt acgtgtctc acttccatgc tcacccgac
acggcgaga cggctaagcg taggtgtggc agggatctc cccctctt
tcagctagcc agctgtctc gccttcctt aaggcaacat gcactacc
ccggacgtt acctcatcg gccaacccctt ctgtggccgc aggagatgg
accccggtgg agtcagaaaa taaggtagt attttgact cttcgagcc
gaggaggatg agggaaatgt atccgttcccg gggagatcc tgccggagg
cctcgacgca tgcccatatg gacccggccg gattacaacc ctcacttgtt
aaggaccgg actacgtccc ttccatgtt caccgtgtc cattggcc
cctcgatata cacccacg gaggaaaggagg acgggtgtc tgtaga
tctgccttgg cggagctcgc cacaaggacc ttccgttccgc acggc
agccgcacgg caacggctc ttctgaccag ccctccgc acggc
gttagtgcgtt actccatccat gccccctt ggggggg
gacgggtctt ggttaccgt aacggaggag gctagtggg acgtcg
tcctacacat ggacggcgc cctgtatcacc ccatcgctg cggaggaa
atcaatgcac tgacactt ttgtccgtt caccacact tggctatgc
cgccgcgaa gcctcgccg gaagaaagggtt accttgcaca gactcg
caactccggg acgtgtcaaa ggagatggg gcaaggcg
ctatccgtgg aggaaggctgtt taagctgttcc ccccccatt
tatggggcaaa aggacgtccg gaaacctatcc agcaaggccg
tggaaaggact tgcttggaaa cactgagaca ccaatttgaca
gaggtttctt ggttccaaacc agagaagggg gggcccaagc
ccagattttt ggttccgtt gtgcgagaaa atggccctt acgtgtgt
cctcaggccg tgatgggttcc ttccatcgatc ttccatact
ttccctgttga atgcctggaa agcgaagaaa tgccctatgg
tggttgcact caacggtcac tgagaatgtac atccgttgg
tggtactgg ccccccaaggc cagacaggcc ataagg
ggggccccc tgactaattt taaaggccg aactcggtt atcgcc
gggtactga cgaccagctg cggtaatacc ctcacatgtt
tgtcgagctg cgaagctcca ggactcgacg atgtctgtt
atctgtaaa gcgccggac ccaaggaggac gagc
atgacttagat actctgcccc ccctggggac ccggccaaac
ataacatcat gtcctccaa tggatcgatc ggcacatgt
tatctcccccgtt gtcacccac caccctt
actccatcgta attctggctt aggaacatc atcatgtat
atgatectt tgacttccattt ctcttagtgc
cttagtgc ttagtacgg ggctgttac
attcaacgac tccatggctt tagcgtt
aatagggtgg cttcatgtt cagggaaactt
cgggccagaa gtgtccgc
gggttccgtt
ggggccccc
gggtactga
tgacttacgg
ggctgttac
attcaacgac
aatagggtgg
cgggccagaa

aagtacctct tcaactgggc agtaaggaccc aagctaaac tcactccaat cccggctg 7620
tcccagggttgg atttatccag ctgggtcggtt gctggttaca gcggggggaga catatacac 7680
a^cc^tgtc^ttc^t gtggccggacc ccgctgggtt atgtgggtgcc tactcctact ttctgttaggg 7740
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaaatagg 7800
ccatccgtt tttttccctt ttttttttcc ttttttttcc ttttttttcc ttttttttcc 7860
ttttccctt tttttccctt ctttttttcc ttttttttcc ttttttttcc ttttttttcc 7920
cttagtcacg gcttagtcgtg aaagggtccgt gagcgcgttg actgcagaga gtgtgtatac 7980
tggcctctctc gcaagatcaag 8001

<210> 4
<211> 7989
<212> DNA
<213> Hepatitis C Virus

<400> 4
ggcagcccc gattggggc gacactccac catagatcac tccctgtga ggaactactg 60
tcttcacgca gaaagcgct agccatggcg ttagtatgag tgctgtgcag cttccaggac 120
ccccctccc gggagagcca tagtgtctg cgaaaccgg gatcacacgg gaattccag 180
gacgaccggg tccttctt gatcaaccc ctaatgcct ggatgttgg gcgtgcggcc 240
gcgagactgc tagccgatg gtgtgggtc gchaaaggcc ttgtgttgc gcctgtatagg 300
gtgtgtcga gtgcggggg aggtctcgta gaccgtgcac catgacgacg aatccaaatc 360
ctcaaaaaaa aaccaaaggc cgcccatga ttgaacaaga tggattgcac gcaggttctc 420
cgcccgctt ggtggagagg ctattcgct atgactggc acaacagaca atcgctgtct 480
ctgatgccgc cggttccgg ctgtcagcgc agggggcggc gttttttt gtcagaccc 540
acctgtccgg tgcctgaat gaactgcagg acgaggcagc gggctatcg tggctggca 600
cgacggcgt tccttgcga gctgtctcg acgttgcac tgaagcggga agggactggc 660
tgctattggg cgaagtgcgg gggcaggatc tcctgtcata tcacccgtt cctggccgaga 720
aagtatccat catggctgt gcaatgcggc gggtgcatac gctgtatccg gctacactgc 780
cattcgacca ccaagcgaaa catcgatcg akgcagacg tactcgatg gaagccgggt 840
ttgtcgatca ggtatctg gcaagaaatc atcaggggct cgcgcagcc aactgttgc 900
ccaggtcaaa ggccgcgcatg cccgacggcg aggatctcg ctgtacccat ggcgatgcct 960
gcttgcgaa tatcatggtg gaaaatggcc gttttctgg attcatcgac tggccggc 1020
tgggtgtggc ggaccgtat caggacatag cggtgttgc acgtgtatatt gctgaagagc 1080
ttggcgccga atgggtgcac cgcttcctcg tgctttacgg tatccggcgt cccgattcgc 1140
agcgatcgc ctctatcg ctcttgcac agttctctg agttaaaca gaccacaaatc 1200
gttccctctt agcggatca atccggcccc tcccccctcc ccccccctaa ctttttttgc 1260
cgaagcgc tggaaataagg ccgggtgtcg tttgtetata tggatatttgc caccatattg 1320
cggttttttgc caaatgtgag ggccggggaa cctggccctg tcttgcac gacgttgc 1380
agggtcttttgc cccctctcgc caaaggaaatg caagggtctgt tgaatgtcgt gaaggaagca 1440
gttccctctt aagttcttgc aagacaaaca acgtctgtag cgaccctttg cagggcagcgg 1500
aaccctccac ctggcgacag gtgcctctgc gccaatggc cacgtgtata agatacacct 1560
gcaaaaggcgg cacaacccca gtgcacgtt gtgagttga tagttgtga aagaatcaaa 1620
tggctctctt caagcgattt caacaagggg ctgaaggatg cccagaaggt accccatttg 1680
atgggatctg atctggggcc tgggtgcaca tgctttacat gtgttttagt gaggttaaaa 1740
aacgtcttagg ccccccggaaac cccggggacg tgggtttctt ttggaaaaca cgataataacc 1800
atggcgccca ttacggctt cttccaaacag acggcaggcc tacttggctt catcatcaact 1860
agcctcacag gccggacacg gaaaccaggc gaggggggg tccaaatgtgttgc 1920
acacaatctt tcctggcgac ctgcgtcaat ggctgtgtt ggactgtcta tcatgggtcc 1980
ggctcaaaga cccttgcgg cccaaaggcc ccaatcaccc aaatgtacac caatgtggac 2040
caggacctcg tcggctggca agcccccggg gggcgcgtt cttgacacc atgcacactc 2100
ggcagctcg accttactt ggtcacgagg catgcgtatg tccatccgt ggcggccggc 2160
ggcgacaga gggggggcc actctcccccc agggccgtct ctaatgtgaa gggctcttcg 2220
ggcgggtccac tgctctggcc ctcggggcact gctgtggcata tcttgcggc tgccgtgtgc 2280
accggggggg ttgcgaaggc gttggactt gtaaccctcg agtctatgg aaccactatg 2340
cggtttttttcc ttttcacggg caactctgtcc cttccggccg taccgcagat atccagggt 2400
gccccatctac acggccctac tgtagcgcc aagagacta aggtccggc tgctgtatgca 2460
gccccagggt ataagggtct tgctctgaac cctgtccgtcg cccgcaccct aggtttcggg 2520
gcgtatatgt ctaaggcaca tggatcgac cctaacatca gaaccgggt aaggaccatc 2580
accacgggtg ccccatcac gtactccacc tatggcaagt ttctggccga cgggtgttgc 2640
tctggggggc cctatgacat cataatatgt gatgagtgcc actcaactga ctgcaccact 2700
atctctggca tcggcacat cctggaccac gggggacgg ctggagcgcg actctgtgt 2760
ctcgccaccc ctacggctcc ggatcggtc accgtgcac atccaaacat cgaggagggt 2820
gtctgtcga gcaactggag aatccccctt tatggcaag cccatccccat cgagaccatc 2880
aagggggggg ggcacccat tttctggccat tccaaagaaatgatgtga gtcgcggcc 2940
aagctgtccg gcctcgact caatgtgtg gcatattacc gggcccttga tggatccgt 3000
ataccaacta gcccggacgt cattgtcgta gcaacggacg ctctaatgac gggcttacc 3060
ggcgatttcg actcagtgt cgaactgcac acatgtgtca cccagacagt cgacttcage 3120
ctggaccoga ctttcacccat tgagacgacg accgtgcac aagacgcggt gtcacgtctc 3180
cagccggcgg gcaaggactgg tagggggcagg atgggcattt acaggttgc gactccagg 3240
gaaacggccct cgggcacatgg cgttccctcg gttctgtcg agtctatgac cgccggctgt 3300
gttgtgtac agtcacgcg cccggagacc tcgttggat tgcgggtta cttaaacaca 3360
ccaaagggttcg cctgtccca ggacccatgt gagggtctgg agagcgttgc tagaggcctc 3420
ccccacatag acggccattt ctgtcccaactaagcaggcagg cttccacccatc atcggtggac 3480
ttgttagcat accaggctac ggtgtgcgc agggctcagg ctccacccatc atcggtggac 3540
caaataatgtggaa agtgcacatc acggctaaag cctacgtcgc acggccaaac gcccctgtctc 3600

ttatccagct ggtcggtgc tggttacagc gggggagaca tatatcacag cctgtcttgt 7680
gcccccccccc gctgggtcat gtgggtcccta ctccctactt ctgtagggtt aggcatctat 7740
ctactcccca accgatgaac ggggagctaa acactccagg ccaataggcc atccctgtttt 7800
tttccctttt tttttttctt tttttttttt tttttttttt ttctccctttt 7860
tttttcctct ttttttcctt ttcttcctt tggtggctcc atcttaagccc tagtcacggc 7920
tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc 7980
agatcaagt 7989

<210> 5
<211> 8649
<212> DNA
<213> Hepatitis C Virus

<400> 5
gccccccccc gattggggc gacactccac catacatcac tccccctgtga ggaactactg 60
tcttcacgca gaaagcgct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccccccc gggagagcca tagtggctcg cgaaaccgg gagaacaccg gaattggccag 180
gacgaccggg tccttcttg gatcaaccgg ctcaatgcct ggagatttg gggtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggact gcctgatagg 300
gtgcttgcg gttccccggg aggtctcgac gaccgtgcac catggacac aatcctaaac 360
ctcaaagaaa aaccaaaccg aacaccaacg ggcgcgccc gatggacaaa gatggattgc 420
acgcaagttc tccggccgt tgggtggaga ggctattcg ctatgactgg gcacaacaga 480
caatcgctg ctctgatgcc gccgttccc ggctgtcage gcaaggggcgc cgggttctt 540
ttgtcaagac cgacctgtcc ggtgcctgaa atgaactgca ggacgaggca ggcggctat 600
cgtggctggc cacgacgggc gttccctgca cagctgtctc cgacgttgc actgaaagcgg 660
gaagggactg gctgttattg ggcgaagtgc cggggcagga tctccgttca tctcaccctt 720
ctccctgcgc gaaatgttcc atcatggctc atgcaatgcg gggctgtcat acgcttgcata 780
cggttactg cccatcgac cacaagcg aacatcgcat cgacgaggca cgtactcgga 840
tggaaaggcg ttttgtcgat caggatgatc tggacaaaa gcatcgaggc ctcgcggcc 900
ccgaaactgtt cgccaggctc aaggcgccgca tgcccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgcgc aatatcatgg tggaaatgg ccgcctttctt ggattcatcg 1020
actgtggccg gctgggtgtg gcccggcgt atcaggacat acgcttggct acccggtata 1080
ttgtgaaga gcttggcgc gaaatggctg accgttctt cgtgtttac ggtatcgcc 1140
ctcccgattt cgcagcgatc gccttctatc gccttcttgc cgagttctc tgagttaaa 1200
cagaccacaa cggttccctc ctacggggat caattccggc cctctccctc cccccccct 1260
aacgttactg gcccggcgc cttggataaa gcccgggtgtg cttttgttca tatgttattt 1320
tccaccat tggcttccctt tggcaatgtg aggccggga aaccgttgc cttttttttt 1380
acgagcattc ctagggtctt tttccctctc gccaaggaa tgcaggatgtc gttgaatgtc 1440
gtgaaggaaag cagttctctt ggaagctt tgaagacaaa caacgtctgt 1500
tgcaggcgc ggaacccccc acctggcgc acgttgcctt gcccggccaaa gccacgtgt 1560
taagatacac ctgcaaaaggc ggcacaaccc cagtggccacg ttgtgatgt 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaaagg tgcccagaag 1680
gtacccatt gtatggatc tgatctgggg ctcgggtca catgttttac atgtgtttag 1740
tcgaggtaa aaaacgtcta ggcggcccgca accacgggg cttgttttc ctttggaaaa 1800
cagataata ccatggaccc gggatggca gcatcgctg gggcgccgt ttcgttaggt 1860
ctgatactt tgacttgc accgcataat aagcttgc tgcgttgc ttacaatatt 1920
tttaccat tttaccatg ggcggaggca cacttgcag ttttttttcc cccctcaac 1980
gttcgggggg gcccgatgc cgtcatccctc ctcacgtgc cgatccaccc agagctaattc 2040
tttaccatca ccaaaatctt gtcggccata ctcggccac tcatgtgtc ccaggctgtt 2100
ataaccaaag tggctactt cgtcgccgc cacgggtca ttcgtgcgtg catgctgtg 2160
cggaagggtt ctgggggtca ttatgtccaa atggcttca tggatgttgc cgcactgaca 2220
ggtagtacg tttatgtacca ttcacccccc ctgcggact gggcccaaccc gggcttaacga 2280
gaccttgcgg tggcagggtg gcccggctc tttctgtata tggagacaaa gtttatcacc 2340
tggggggcag aaccccgccg gtgtggggac atcatcttgg gcttgcgttcc cttcccccgc 2400
aggggggggg agatacatct gggacggca gacacgtt gggccggcc 2460
ctcgcccta ttacggctt ctcacccacag acgcggaggcc tacttggctg atgcacccgt 2520
agcctcacag gcccggacag gacccgggtc gggggggagg tccaaatgtt ggcgttgc 2580
acacaatctt tccctggcgac ctgcgtcaat ggcgtgtttt ggactgtctt tcatgttgc 2640
ggctcaaaaga cccttgcggg cccaaaggcc ccaatcaccc aaatgtacac caatgtggac 2700
caggaccccg tccggctggc acgttactt gtcacggagg catggcgtatc tcatgttgc 2760
ggcagctgg accttactt gtcacggagg catggcgtatc tcatgttgc ggcggccggg 2820
ggcgcacacg gggggggccctt acttccccc agggccgtt cttatgttgg gggcttccgt 2880
ggcggccac tgccttgcctt ctcggggcac gtcgtggca ttttcggc tgcgtgtgc 2940
acccgggggg ttgcgaaggc gttggacttt gtaccgtcg agtctatggaa aaccactatg 3000
cggttccccc tcttcacggc caactcgatc ctcggcccg taccggacac attccagggt 3060
gccccatctac acgccttac tggtagggc aagagcacta aggtggccggc tgcgtatgca 3120
gccccaaagggtt ataagggtct tttctgtacac cctgcgttgc cccggccaccc aggtttcg 3180
gctgtatgtt ctaaggcaca tggatgtcgtac cttacatca gacccggggt aaggaccatc 3240
accacggggt ccccccacat gtaactccac ttttgcgttgc ttttgcgttgc 3300
tctggggggc cttatgtacat cttatgtt gatggatgtcc acttcactgaa ctcgaccatc 3360
atctctggcc tccggcactgt cttggaccaaa gcccggacgg ctggagccgcg actcgtcg 3420
ctcgccaccc ctacggccctt cggatgggtc accgtggccac atccaaacat ggggggggt 3480
gctgttgcctt gactggaga aatccctt tttctgtacat gatggacaaa ccatccccat cgagaccatc 3540
aagggggggg ggcacccat tttctgtacat tccaaaaga aatgtatgt gtcggccgc 3600

aagctgtccg gcctcggact caatgctgt acaatttacc gggccttga tgcgttcgtc 3660
 ataccaacta gcgaggacgt cattgtcgta gcaacggacg ctctaattgac gggctttacc 3720
 ggcatgtcg actcaactgt cgactgtca acatgtgtca cccagacagt cgacttcagc 3780
 ctggaccgcg ccttcaccaat tgagacgacg accgtgccac aagacgcggg gtcacgtcg 3840
 cagcggcgg gcaaggactgg tagggcagg atggcattt acaggtttgt gactccagga 3900
 gaacggccct cgggcatgtt cgattcctcg gttctgtcg agtgctatga cgccggctgt 3960
 gcttggtagc agtcacgac cggcggaccc tcagtttagt tgcgggctta cctaaacaca 4020
 ccagggttc cggctctgcca ggaccatctg gagttctgg agagcgttctt tacaaggcctc 4080
 acccacatag acgcccattt cttgtccccag actaaggcagg caggagacaa cttcccttac 4140
 ctggtagcat accaggctac ggtgtgcggc agggtctagg cttcacccctc atcggtggac 4200
 caaatgttga agtgcgtcat acggctaaag cttacgtcg acggggccaa gcccctgtcg 4260
 tataggctgg gagcgttca aacagggtt acataccacac accccataac caaatatcatc 4320
 atggcatca tgcgtctgt cctggaggc gtacacgac cctgggtgtt ggtaggcgga 4380
 gtcctagcg ctctggccg gttttccctg acaacaggca gctgtgtcat tgcgtggcagg 4440
 atcatctgt ccggaaagcc ggcacatcatt cccgacaggg aagtccctta cccggagttc 4500
 gatgagatgg aagagtgcgc ctcacaccc cttacatcg aacagggaaat gcaagctcgcc 4560
 gaacaatca aacagaaggc aatcggttg ctgcaaacag ccaccaagca agcggaggct 4620
 gctgctcccg tgggttcaatc aagtgccgg accctcgaaag ccttctggc gaagcatatg 4680
 tggaaatttca tcagcggggat acaatattttt gcaaggcttgc ccaactcgcc 4740
 gcgatagtc cactgtatgc attacagcc tttatccca gcccgttcac caccacat 4800
 accctctgt ttaacatctt gggggatgg gtggccggcc aacttgcgttcc tcccaggcgt 4860
 gtttgcgtt tcgtaggcgc cggcatcgct ggagcggctg ttgcagcat aaggcttgg 4920
 aagggtctt tggatattttt ggcaggattt ggagcagggg tggcaggcgc gtcgtggcc 4980
 tttaaaggta tgagcggcga gatgcctcc accgaggacc tggtaacatc acgtccctgt 5040
 atccctcccc ctggccccc agtgcgtggg gtcgtgtcg cagcgataact gcaagtcggcac 5100
 gtggcccg gggaggggc tgcgtgggg atgacccgc tgcgtgcgtt cgcttcggcg 5160
 ggttaaccacg tctcccccac gcaactatgt cctgagacgc acgctgcaggc 5220
 cagatccct ctatgtttac catcactcg tgcgtcaaga gcttcaccaatc gttgatatgc 5280
 gaggactct ccacccatcg ctccggctcg tggtaaagag atggttgggaa 5340
 acgggttga ctgatttcaa gacctggctc cagtcacccgc tccgtccggc attgcggga 5400
 gtccctcttct tctcatgtca acgtgggtac aaggagatct ggcggggcga cggcatcatg 5460
 caaaaccatc gcccacatgtt gacacagatc accggacatcg tggatccatc 5520
 atcggtggcc ctggacatcg tagtaacacg tggcatggaa cattccctat taacgcgtac 5580
 accacggcc cctgcacgc cttcccccggcg ccaaaattttt ctggggcgat gtcgggggt 5640
 gctgtcgagg agtacgttga gtttacgcgg gttggggatt tccactacgt gacgggcgt 5700
 accactgaca acgttaatgtt cccgtgtca gttccggccc ccaatttccatc 5760
 gatgggttgc gtttgcacag gtacgttca gtcgtcaac cccctcttacg 5820
 acattctgg tcggctcaa tcaataccctt gttgggttac agtcccattt cggagccaa 5880
 cccgacgttgc cagtgcttca cttccatgtctt accgaccctt cccacattac 5940
 gctaaggta ggctggccag gggatctccc cccctcttgg ccagtcatc agctagccag 6000
 ctgtctgcgc cttcccttga ggcacatcg actaccgtc atgactcccc ggaacgttac 6060
 ctcatcgagg ccaaccttctt gtggccggcag gagatggggc ggaacatcac 6120
 tcagaaaata aggttagtat ttttgcgttctt ttcgagccgc tcaagcgga 6180
 agggaaatgtt ccgttccggc ggagatctt cggaggttca gggatatttttccatc 6240
 cccatattgg cccggccggc ttacaacctt ccactgttag agtcttggaa 6300
 tacgtccctc cagtggtaca cgggtgttca ttgcgtccctg ccaaggcccccc 6360
 cttccacggc ggaagaggac gtttgcgttgc tcaagatctt ccgtgtcttc 6420
 gagctcgcca caaaaggatcc cggcagctcc gaatgtcgcc cgtcgacacg 6480
 acggcccttc ctgacccatcg cttccgacacg ggcgacgcgg gatccgacgt 6540
 tccctccatgc ccccccatttga gggggagccg ggggatcccccc attcagcgatc 6600
 ttttccatgc gcgaggaggc tagtgaggac gtcgtctgtc gtcgtatgtc 6660
 acaggccccc tgatcactcgcc atgcgtcgcc gggaaacca agtgccttccatc 6720
 agcaacttctt tgcgtccgttcc cccacatttgc gtcgtatgtca caacatctcg 6780
 ctgcggcaga agaagggttca ctttgcgttca ctgcgttcc tggcggccatc 6840
 gtgtcaatgg agatggatgc gaaaggcttcc acagtttgggg ctttttttttgc 6900
 gaaaggctgttca agtgcacgc cccacatttgc gccagatctt aatttggatc 6960
 gacgtccggc acctatccatcg caaggccgtt aaccacatcc gtcgttccgtt 7020
 ctggaaagaca ctgagacacc aatttgcgttcc accatcatgg caaaaaatgtt 7080
 gtcacaccatc agaagggggg cccgttccatcg gtcgttccatcg tgcgttccatcg 7140
 gtttgcgttgc gtcgttccatcg gtcgttccatcg tgcgttccatcg tgcgttccatcg 7200
 atgggttgc ttttccatgc cccatcgatc cccatcgatc gtcgttccatcg 7260
 gcttccatcg gtcgttccatcg ttttccatgc ttttccatgc ttttccatgc 7320
 acggtcactt gtcgttccatcg ttttccatgc ttttccatgc ttttccatgc 7380
 cccggccatc gacaggccat aagggttccatcg acagggccgc ttttccatgc 7440
 actaatttca aaggccatcg ttttccatgc ttttccatgc ttttccatgc 7500
 accagcttgc gtaataccctt cccatcgatc ttttccatgc ttttccatgc 7560
 aagctccagg actgcacatcg gtcgttccatcg gtcgttccatcg ttttccatgc 7620

gcggggaccc aagaggacga ggcgagccta cgggccttca cgaggactat gactagatac 7680
tctgcccccc ctggggaccc qcccaaaccga aatacactg tggagttgat aacatcatgc 7740
tcctccaaatg tgtagtcgc gcacgatgca tctggcaaaa ggggtgacta tctcacccgt 7800
gaccggcacca ccccccttgc ggcccgtcg tggagacag ctagacacac tccagtcata 7860
tcctggctag gcaacatcat catgtatgcg cccacccgt gggcaaggat gatcctgatg 7920
actcatttct tctccatct tctagctca gaaacaacttg aaaaaggccct agattgtcag 7980
atctacggg cctgttactc cattgagcca cttgacccat ctcagatcat tcaacgactc 8040
catggcccta ggcattttc actccatagt tactctccag gttagatcaa tagggggct 8100
tcatgcctca ggaaacttgg ggtaccgcctt tgtaggtct ggagacatcg ggccagaagt 8160
gtccgcgcta ggctactgtc ccagggggg agggtgcctt ctttgccaa gtacctcttc 8220
aactgggcag taaggaccaa gctcaaactc actccaaatcc cggtgcgtc ccagttggat 8280
ttatccagct ggttgttgc tggtagcagc gggggagaca tatatcacag cctgtctcg 8340
gcccggccctt gctgggtcat gtggtgccctt ctcctacttt ctgttaggggt aggcatctat 8400
ctactcccca accgatgaac ggggagctaa acactccagg ccaataggcc atccctttttt 8460
ttttccctttt tttttttttt tttttttttt tttttttttt ttctccctttt 8520
tttttccctct ttttttctt ttctccctt tggtggctcc atcttagccc tagtcacggc 8580
tagctgtgaa aggtccgtga gccgcttgac tgcagagat gctgatactg gctctctgc 8640
agatcaagt 8649

<210> 6
<211> 8001
<212> DNA
<213> Hepatitis C Virus

<400> 6

ggcagcccc gättggggc gäcaactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgctc ägcatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctccc gggagäggca tagggctcg cggAACGGT gagtacaccg gaattgccag 180
gacgaccggg tccttcttg gatcaacccg ctcaatgcct ggagatttg ggctgcccc 240
gcaagactgc tagccgaäta gtgttgggtc gcaaggcc ttgttgtact gcctgatagg 300
gtgcttgcga gtgcggccgg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaagaaa aaccaaacgt aacccaacg ggccgcgcatt gatggacaaa gatggattgc 420
äcgcagggtt cccggccgt tgjgtggaga ggcttattcg ctatgactgg gacaacacg 480
caatcggtcg ctctgtgcc gccgttccc ggctgtcagc gcaaggggcgc ccggttctt 540
ttgtcaagac cgaacctgtcc ggtgcctgta atgaaactgca ggacgaggca ggcggctat 600
cgtggctggc cacgacgggc gttccttgca cagctgtctc cgacgttgc actgaagcgg 660
gaagggactg gctgttattg ggcaaggatgc cggggcaggat ttcctgtca ttcacccctt 720
ctcctgcga gaaägtatcc atcatggctg atgcaatgcg gcccgtgcac acgcttgcac 780
cggctacccg cccattcgac caccaagcga aacatcgcat cgacgagca cgtactcgga 840
tggaaagccgg tcttgcgtat caggätgtac tggacaaga gcatcaagggg ctcgcggcäg 900
ccgaaactgtt cgcagggtc aäggcgcgc gttcccacgg cggaggatctc gtcgtgacc 960
atggcägtgc ctgttgcgg aatatcatgg tgaaaattgg ccgtttttt ggattcatcg 1020
actgtggccg gctgggtgtg gcccggccgt atcaggacat acgcttggct acccgtgata 1080
ttgtcaagaa gcttggccgc gaatgggtc accgccttctt cgtgtttac ggtatcgccg 1140
ctcccgatcc gcaaggccatc gccttctatc gccttcttgc caagtttctc ttaggtttaaa 1200
cagaccacaa cggtttccctt ctggggat caattccgc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaaataa ggccgggtgtg cgtttgtcta tatgttattt 1320
tccaccatatt tggcgttctt tggcaatgtg agggccggaa aacctggccc tttttttt 1380
acgacgatc ctggggatc ttcccctctc gccaaggaa tggcaaggatc gttgaätgtc 1440
gtgaaggaaag cägttctctt gggaaägttctc tgaaaggaaa caacgttctg agcggccctt 1500
tgcaggccgc ggaacccccc acctggcäac aggtgcctc gcccggccaaa gccacgtgta 1560
taagatacac ctgcaaaaggc ggcacaaaccc cägtggccacg ttgtgatgtg gatagtgtg 1620
gaaagagtc aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatggatc tgatctgggg cctcgtgtca catgttttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggccccccca accacgggg cgtggttttc ctttgaäaaa 1800
cacgataataa ccattggccg tattacggcc tactcccaac agacgcggg cctacttggc 1860
tgcatcatca ctggcctcact aggccgggac aggaacagg tggggggg ggtccaaatgt 1920
gtctccacccg caacaaatcc tttcttgcg acctgcgtca atggcgtgtg ttggactgtc 1980
tätcatgttgc cggcgtcaaa gacccttgcg gcccggccagg gccaatcac ccaaattgtac 2040
accaatgtgg accaggacct cgtcggtctgg caagcggccc cccggggcgcg ttcttgaca 2100
ccatgcaccc gcccggccgtc ggaccttac ttggtcacga ggcacatggccg tgcattccg 2160
gtgcggccgc gggccgacag cagggggagc ctactctcc ceaggccctg ctccctactt 2220
aagggtcttt cggccgggtcc ätgcgtctgc ccctccgggc atgctgtggg cätcatttcgg 2280
gtgcggctgt gcacccgggg ggttgcgaag gcccggact ttgttccctg cgagtctatg 2340
gaaaccacta tgcgttccccc ggttccatc gacaactctg cccctccggc cgtaccggc 2400
acatccagg tggcccatct acacggccccc actggtagcg gcaagagcac taagggtccg 2460
gctcgatgtc cagcccaagg gttaaagggtt cttgttctga acccgccctg cggccggccacc 2520
ctagggttgc gggcgatataa gtcataaggca catggatctg accctaataat cägaaccgg 2580
gttagggacca tcaccacggg tgccccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacgggtgtt gctctgggg cggctatgac atcataataat gtatgtgtg ccactcaact 2700
gactcgacca ctatccctggg catcgccaca gtcctggacc aacgggagac ggctggagcg 2760
cgactcgctg tgctcgccac cgctacggct ccgggatcg tccacgtgc acatccaaac 2820
atcgaggagg tggctctgtc cagcaactggaa aatccctt ttatggcaaa gcccattccc 2880
atcgccacca tcaagggggg gaggccaccc atttttgcg attccaagaa gaaatgtat 2940
gagctccgg cggacatctc cggccctggg ctcatgtcg tagcatatta cccggccct 3000
gatgtatccg tcataccacat tagccggacat gtcattgtcg tagcaacggc cgctctaatt 3060
acgggttta cggccgattt cgactcgatg atcgäctgca atacatgtgt caccacgaca 3120
gtcgacttca gcctggaccc gacccatcacc attgäjacga cgaccgtgc acaagacgc 3180
gtgtcacgt cgcaggccgc aggccggact ggtaggggca ggtatggcat ttacagggtt 3240
gtgactccag gagaacggcc ctggggcatg ttgcatttcc cgggtctgtg cgagtctat 3300
gacggggct gtgttggta cägactcgac cccggccaga cctcaatgtt gttggggct 3360
tacctaaca caccaggctt gcccgtctgc caggaccatc tggatgttgc ggagacgctc 3420
tttacaggcc tcaccacat agacggccat ttcttgcctt agataggca ggcaäggagac 3480
aactccctt acctggatgc ataccaggct acgggtgtcg ccagggtctc gacggccatc 3540
ccatcggtgg accaaatgtg ggatgtctc atacggctaa agcctacgct gacggccatc 3600

acgccccctgc tgtataaggct gggagccgtt caaaacgagg ttactaccac acaccccaata 3660
 accaaataca tcatggcatg catgtcggtc gacctggagg tcgtcacgag caccctgggt 3720
 ctggtaggcg gagtccttagc agctctggcc gcgtattgcc tgacaacagg cagcgtggc 3780
 atttgtggca ggatcatctt gtcggaaag cggccatca ttcccggacag ggaagtccct 3840
 taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcctttacat cgaacaggaa 3900
 atgcagctcg ccgaaaaatt caaacagaag gcaatcggt tgctgcaaa agccaccaag 3960
 caagcgagg ctgtcgctcc ctgtgtggaa tccaatggc ggacccctga agcccttgg 4020
 gcaagcata tgtggatatt catcaacggg atacaatatt tagcaggctt gtccacttg 4080
 cctggcaacc ccgcatacg atactgtat gcattcacag cctctatcac cagccgcctc 4140
 accacccaaac ataccctctt gttiacatc ctgggggat gggggccgc ccaacttgct 4200
 cctcccaacg ctgtttctgc ttctgttaggc gccggcatcg ctggagcgcc tttggcagc 4260
 ataggccctg ggaagggtgt tttggatatt ttggcagggtt atggagcagg ggtggcaggc 4320
 gcgctcgcc cctttaagggt ctatggcgcc gagatgcctt ccacccggaga cctgggttaac 4380
 ctactccctg ctatcccttc ccctggccgc ctatgtcgcc gggtcgctgtg cgcaagcgata 4440
 ctgcgtcgcc acgtggggcc agggtgggg gctgtcgatg ggtgaaaccc gctgatagc 4500
 ttctgttcgc ggggttacca ctgtttcccccc acggactatq tttgtggagag cgacgctgca 4560
 gacacgtgtca ctcagatccct ctctagttt accatcaactc agtgcgtgaa gaggcttcac 4620
 cagtggatca acggggactg ctccacgccta tgctccggct ctgtggctaaag agatgtttgg 4680
 gattggatat gcacgggttt gactgattt aagacctggc tccagttcaa gtcctggccg 4740
 cggattggccg gagtccctt ttctctatgt caacgtgggtt acggggatggat gttggccggc 4800
 gacacgtgtca tgcaaaaaccac ctggccatgt ggagcacaca tcacccggaca tttggaaaaac 4860
 ggttccatga ggatcggtgg gccttaggacc tttgttttttca cgtggcatgg aacattcccc 4920
 attaacccgtt acaccacggg ccctgtcgcc ccctccccgg cggccaaatttttcttagggcg 4980
 ctgtggccgg tggctgtctg ggttacgtt ggggttacgc ggggtggggat tttccactac 5040
 gtgacgggca tgaccactga caacgttttttca ttttttttttca aggttccggc ccccgaaattt 5100
 ttcacagaaag tggatgggg tgggttgcac aggtacgctc cagcgtgcac acccctctta 5160
 cggggaggagg tcacatttccctt ggtcggtcctt aatcaatacc tgggtgggtc acagctccca 5220
 tgccggcccg aacccggacgt acgtgtctc acttccatgc tcacccggacc cttcccaatattt 5280
 acggccggaga cggcttaacgg taggctggcc agggggatctc cccctcttccctt gggccagctca 5340
 tcaactgtacc acgtgtctgc gccttcctt aaggcaacat gcactaccgg tcatgactcc 5400
 cgggacgttcc acctccatcgaa ggccaaacccctt ctgtggccgc agggatgggg cgggaaacatc 5460
 accccgttgg agtcaaaaaaa taaggtagta atttttttttctt ctttcgtatcc gtccttccat 5520
 gaggaggatg agaggaaatg atccgttcccg gggggatcc tggggggatcc cggggaaattt 5580
 cttccgtatcc tggggggatcc ttttttttttctt cttccactgtt gtttttttttctt cttccactgtt 5640
 aaggacccttcc actacgttcc tccagtggta cacgggtgtc cattggccgc tgccaaaggcc 5700
 cttccgtatcc caccctccacg gaggggaggg acgggttgc ttttttttttctt cttccactgtt 5760
 tctgccttgg cggagctcgcc cacaaggacc ttccggcagctt ccgaatcgcc gggccgtcgac 5820
 acggggccacgg caacccgttcc ttctgtaccgg acggccggcc acggccggcc gggatccggac 5880
 gttggatgtt acttccatccat gggggggatcc ttttttttttctt cttccactgtt cttccactgtt 5940
 gacgggttcc ggttccatccat gggggggatcc ttttttttttctt cttccactgtt cttccactgtt 6000
 tccatccatccat gggggggatcc ttttttttttctt cttccactgtt cttccactgtt cttccactgtt 6060
 gtcaatgcac ttttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 6120
 cggccggccaa gcttgcggca gaagaaggcc accttttgcac gactgcagggtt cttccactgtt 6180
 cactacccttcc acgtgtctca gggggggatcc ttttttttttctt cttccactgtt cttccactgtt 6240
 ctatccgttgg aggaagccctg taaggctgtacg ccccccacattt cggccagatc taaattttggc 6300
 tatggggccaa aggacgttccg gaaaccttatcc agcaaggccg ttaaccacat ccgtccgtt 6360
 tggggggccaa aggacgttccg gaaaccttatcc agcaaggccg ttaaccacat ccgtccgtt 6420
 ggggttttttctt ggttccatccat gggggggatcc ttttttttttctt cttccactgtt cttccactgtt 6480
 ccagatttgg ggggttccatccat gggggggatcc ttttttttttctt cttccactgtt cttccactgtt 6540
 cttccactgtt cttccactgtt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 6600
 ttccctggta atggctggaa agcgaagaaa tggccatgtt gcttcgtatcc ttttttttttctt cttccactgtt 6660
 ttttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 6720
 ttttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 6780
 gggggccccc ttttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 6840
 ggttgcgtt cttccactgtt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 6900
 ttttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 6960
 atctgtggaa gggggggccaa ccaaggaggac gggggggccaa ttttttttttctt cttccactgtt 7020
 atgacttagat acttccatccat cccctggggatcc ttttttttttctt cttccactgtt cttccactgtt 7080
 ataacatcat gtttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 7140
 tatctccatccat gtttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 7200
 acttccatccat gtttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 7260
 atgatccatccat gtttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 7320
 ctatccatccat gtttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 7380
 attcaacccatccat gtttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 7440
 aatagggttttccatccat gtttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 7500
 cggggccaggaa gtttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 7560
 aagtaccttccatccat gtttttttttctt cttccactgtt cttccactgtt cttccactgtt cttccactgtt 7620

tcccàgttgg atttatccag ctggttcggt gctgggtaca gcgggggaga cátatatcac 7680
ágccctgtctc gtgcccgácc ccgcgtggtc atgtggtgcc tactcctact ttctgttaggg 7740
gtäggcätct atctactccc caaccgatga acggggägct aaacactcca ggccaatagg 7800
ccatccctgtt tttttccctt ttttttttc tttttttttt tttttttttt tttttttttt 7860
ttttctccctt tttttttctt cttttttcc ttttctttcc ttgggtggct ccatcttagc 7920
cctagtacag gctägctgtg aaaggccctg gagccgctt actgcagaga gtgctgatac 7980
tggccctctt gcagatcaag t 8001

<210> 7
<211> 11076
<212> DNA
<213> Hepatitis C Virus

<400> 7

```

gcccggcccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctccccc gggagagcca tagtggtctg cggAACCGGT gagtacaccg gaattgccag 180
gacgaccggg tccttcttg gatcaacccg ctcaatgcct ggagatTTGG gcgtgcgggg 240
gcgagactgc tagccgaga gtgttgggtc gcgAAAGGC ttgtggtaact gcctgatagg 300
gtgcttgcgatgtggccgg aggctctcgat gacgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aaccaacacg ggccgcacat gattgaacaa gatggattgc 420
acgcaggatcc tccggccgt tgggtggaga ggtatcgg ctatgactgc gcacaacaga 480
caatcggctg ctctgatgcc gcccgtttc ggtgtcagc gcaaggggcgc cgggttttt 540
ttgtcaagac cgacctgtcc ggtgcctctg atgaactgcg ggacggaggca ggcgcgctat 600
cgtggctggc cacgacgggc gttccctgctc cagctgtctc cgacgttgc actgaadcg 660
gaagggactg gctgttattt ggcgaagtgc cggggcaggat tctcctgtca ttcacccctt 720
tcctggccgaa gaaagtatcc atcatggctg atgcaatgcg gcccgtgcac acgcttgc 780
cggttacccg cccattcgac ccccaagcga aacatcgat cggcggcggca cgtactcg 840
tggaaaggccg tcttgcgtat caggatgtatc tggacaaga gcatcagggg ctcgcggcc 900
ccgaaactgtt cggcaggcgc aaggcgcgc tggccggcgg cgaggatctc gtcgtgaccc 960
atggcgatgc tggcgtccg aatatcatgg tgaaaatgg ccgttttttgc gattcatcg 1020
actgtggccg gttgggtgtg gggggccgtt atcaggacat acgcttgcgt accctgtata 1080
ttgtcaagaa gcttggccgc gaatgggctg accgttctt cgtgttttac ggtatcgcc 1140
ctcccgatcc gcagcgcatac gccttctatc gccttcttgc ctagttcttgc tgatttaaa 1200
cagaccacaa cgggttccctt ctagcggat caattccgc cctctccctc cccccccccc 1260
aacgttactg gccgaaggccg cttggataaa ggccgtgtg cgttgcgtca tatgttattt 1320
tccaccatattt tggcaatgtg agggccggaa aacctggccc tggcttcttgc 1380
acgaggatcc ctggggctt tttccctctt gccaaaggaa tgcgaaggatc gttgaatgtc 1440
gtgaaggaaatgtt cgttccctt ggaagatctt tgaagacaaa caacgtctgt aacgcaccctt 1500
tgcaggcggc gggaaaaacc acctggcgcac aggttgcctt gggccaaaaa gccacgtgt 1560
taagatCACAC ctggccaaaggc ggcacaaccc ctaggtccacg ttgtggatgg gatagtgt 1620
gaaagatgtca aatggctctc tcaacggatc ttcaacaagg ggctgtggaa tgcccagaag 1680
gtacccattt gtatggatc tgatctgggg cctcggtca catgttttac atgtgtttag 1740
tcgaggatataaaaacgtctca gggcccccgg accacggggg cttgtgttttgc tttgaaaaaa 1800
cagataata ccatgggcac gaatccataaa cctcaaaagaa aaaccaaacg taacaccaac 1860
cgccggccac aggacgtcaa gttccggggc ggtggcaga tggcgtgg agtttaccc 1920
ttggccgcga gggggcccccgg gttgggtgtg cggcggacta ggaagacttc cgagcggctg 1980
caacctcgatc gaaaggcgcata acatcccccc aaggctcgcc agcccgagg tagggccctgg 2040
gctcggcccg ggttcccttgc gggccctatc ggcattgggg gtttgggggttgc ggcaggatgg 2100
ctctgttac cccgtggctc tggcgttgc tggggccca cggacccccc cgttgcgttgc 2160
cgcaattttt gtaaggatc ctagtacccctt acgtggggct tggccgtatc catgggttac 2220
attccgtcg tcggccggcc ctttggggc gtttgcgggg ccctggccca tggcgtccgg 2280
gttctggagg acggcgtgaa ctatgcacca gggaaatctgc cgggttgc ttttctatc 2340
ttcttcttgc ttttgcgttgc ctgtttgcacc atcccaagttt cccgttatgc agtgcgcac 2400
gtatccggatgtt gtttccatgt cagcggatc tggccaaacgc caacgttgcgt gttatggc 2460
cgggacatcg tcatcatacc ccccccgggtgc gttccctgc ttcggggagaa caactccctcc 2520
cgctgtggatgttgc tccacgttgc tggccggaggac acgttagctt cccactacg 2580
acgatcgatc gccatgtcgat tttgtcgat gggggccggcgttgc ctctgttgc cgtatgtac 2640
gtggggatgc tctggggatc ttttttgc tggccggatc gtttgcgttgc ttttctatc 2700
cgccacgaga cgtacaggatc tggcaattgc tcaatatacc cggccacatgt gacagggttac 2760
cgatggatcc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 2820
ttactccggatccaccaacgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 2880
ggccggcccttgc ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 2940
ctcttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3000
ctggggatgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3060
aaccggcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3120
tttcccttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3180
ggccggcccttgc ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3240
tttcccttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3300
tttcccttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3360
ggccggcccttgc ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3420
tttcccttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3480
tttcccttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3540
tttcccttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc 3600

```

tgggttcgg ggccttggtt gacacccaga tgcttggtcc actaccata caggcttgg 3660
 cactaccct gcactgtcaa cttaaccatc ttcaaggtta ggatgtacgt ggggggagtg 3720
 gagcacaggc tcgaagccgc atgcaatttg actcgaggag agcgttgaa cttggaggac 3780
 aggacagat cagacttag cccgtctgtc ctgtctacaa cggagtggca ggtattgccc 3840
 tggcttcggcc ccacccttacc ggctctgtc actgggttgc tccatctcca tcagaaacgtc 3900
 gtggacgtac aataacctgtc cggtataggg tggcggttg tctcccttgc aatcaaatgg 3960
 gagtatgtcc tgggtctt ccttcttgc gggacgcgc ggtctgtgc ctgcttgg 4020
 atgatgtgc tgatagctca agctgaggcc gcccstagaga acctgggtt cctcaacgcg 4080
 gcatccgtgg ccggggcgca tggcatttcc tccttcttgc tggcttgc tggcttgc 4140
 tacatcaagg gcaggctgtt ccctggggcg gcatatgcgc tcaacggcgt atggccgcta 4200
 ctccctgtcc tggctgggtt accaccacca gcatacggca tggaccggga gatggcaigca 4260
 tggcggtgg ggcgggttt ctgtggctt atacttgc caatatttca tccacaggcc gcaactaaag 4320
 ctgttcttcg ctaggctcat atgggttgc caatatttca tccacaggcc cgaggccac 4380
 ttgcaagtgt ggatcccccc cctcaacgtt cggggggggc gcatatgcgtt catcccttc 4440
 acgtgcgcga tccacccaga gctaattttt accatcacca aatatttgc cgccatactc 4500
 ggtccactca tgggtctca ggctgttata accaaatgtc cgtaacttcgt ggcgcacac 4560
 gggctcattc gtgcattgc gctgggtccgg aagggttgc ggggttattt tggccaaatg 4620
 gctctcatga agtggccgc actgacaggt acgtacgtttt atgaccatct caccctactg 4680
 cgggacttgg cccacggggc cctacagac ctggcggtgg cagttgagcc cgtcgcttc 4740
 tctgtatgg agaccaaggtt tatacttgc gggcagaca cccggcggtg tgggacatc 4800
 atcttggcc tggccgttc cggccggcgg gggaggaga tatacttgc accggcagac 4860
 agccttggaa ggcagggggtt ggcacttc ggccttata cggcctactc ccaacaaag 4920
 cgaggcttac ttggctgc tatacttgc ctcacaggcc gggacaggaa ccaaggctcag 4980
 ggggagggtcc aagtggtctc caccgcaaca caattttcc tggcgttgc gtcattttgc 5040
 gtgtgttggaa ctgtctatca tgggtccggc tcaaaggaccc ttggccggccc aaaggccca 5100
 atcaccctaaat tgacaccatc tgggaccatc gacccatgc gctggcaagc gccccccggg 5160
 ggcgttctt tggccatgc cacccatgc gacccatgc tttacttgc gacggatc 5220
 gccgtatca tttccgttgc cggccggggc gacccatgc gggccatctc cttcccttc 5280
 cccgtcttctt aacttggaaat ctccctggc ggtccatgc tctcccttc gggcatgt 5340
 gtgggcatct ttcgggttgc ctgtgttgc cgggggttgc cggaggccgtt gggactttgt 5400
 cccgtcgatgt ctatggaaac cactatgcgg tcccttgc tcaacggacaa ctcgtccct 5460
 cccggcgatc cggcagacatt ccagggtggc catctacacg cccctactgg tagcgccaaag 5520
 agcaactaagg tggccgttgc gatgtcggcc caagggtata aggttgc tttttttttt 5580
 tccgtcgccg ccacccttgc ttccggggc tatatgttca aggcacatgg tatacgacc 5640
 aacatcagaat ccggggtagg gacccatcacc acgggttgc ccatcagtc tccacccat 5700
 ggcaagttt tggccgttgc tgggttgc tggggggccgtt atgacatcat aatatgtgt 5760
 gaggccacta caactgactc gacccatc tgggcatec gacagtc tttttttttt 5820
 gagacggctg gagccgtact cgtcggttgc gccacgtt cccctccggg atcggttacc 5880
 gtgccacatc caaacatgc ggggttgc tggccacgc ctggagaaat ccccttttgc 5940
 ggcaaaagcca tcccatcgc gaccatcaag gggggggggc accttattttt ctggccat 6000
 aagaagaat gtgtgttgc cggccgttgc tggccgttgc tggacttca tggccat 6060
 tattaccggg gccttgc tttccatgc gatgttgc tttttttttt 6120
 acggacgttc taatgtacggg ctttaccggc gatgttgc tttttttttt 6180
 tgggttgc accatgc tttccatgc gacccatgc tttttttttt 6240
 gtggccatc accgttgc tttccatgc gacccatgc cggccggcc gggactttttt 6300
 ggcattttac ggttttgc tccaggagaa cggcccttcgg gcatgttgc tttttttttt 6360
 ctgtgttgc gctatgttgc gggctgttgc tggtacggc tttttttttt 6420
 gtttaggttgc gggcttaccc aacacacca gggcttgc tttttttttt 6480
 ttcttggaga gcttgc ttttac accatgc tttttttttt 6540
 aggccaggcc gggccatc accatgc tttttttttt 6600
 gctcagggttc cacccatgc gtttgc tttttttttt 6660
 acgttgcaccc ggcacccatc ctttgc tttttttttt 6720
 accacacacc ccataaccatc atatcatgc gcatgttgc tttttttttt 6780
 acggacatcc ggggttgc tttttttttt 6840
 acaggccatc ttttgc tttttttttt 6900
 gagacggaaat ttttgc tttttttttt 6960
 tacatcgttaccc accatgc tttttttttt 7020
 caaaacggcc ccaacgttgc tttttttttt 7080
 ctgtgttgc ttttgc tttttttttt 7140
 gggccatc ttttgc tttttttttt 7200
 atccatcacc ccaacatc ttttgc tttttttttt 7260
 gggccatc ttttgc tttttttttt 7320
 gggccatc ttttgc tttttttttt 7380
 gggccatc ttttgc tttttttttt 7440
 gggccatc ttttgc tttttttttt 7500
 gtgtgttgc ttttgc tttttttttt 7560
 aaccggctga tagcgttgc ttttgc tttttttttt 7620

gagagcgacgc ctgcagcaca tgtaactcag atccctctca gtcttaccat cactcaactg 7680
ctgaaggaggc ttaccaggc gatcaacgag gactgctcca cgccatgctc cggctcggt 7740
ctaagagatg tttggattt gatatgcacg gtgttgactg atttcaagac ctggctccag 7800
tccaagctcc tgccgcatt gccgggagtc cccttcttcat ctagtcaacg tgggtacaag 7860
ggagtctggc ggggcacgg catcatgcaa accacctgcc catgtggac acagatcac 7920
ggacatgtga aaaacggtc catgaggata gtggggccta ggacctgtag taacacgtgg 7980
catggacat tccccattaa cgcgtacacc acggggccct gcacgcctc cccggcgcca 8040
aattattcta ggccgttgcat gcccggat gctgaggat acgtggaggat taacgtgg 8100
ggggatttcc actacgtac gggcgtacc actgacaaacg taaaatggcc tggtcagggt 8160
ccggcccccg aattttcac agaagtggat ggggtcggt tgacacaggta cgctccacg 8220
tgcaaaaaaa tcctacgggaa ggaggtcaca ttccctgtcg ggctcaatca atacctgggt 8280
gggtcacagc tccatgcgg gcccgaaccg gacgtacag tgctacttc catgctcacc 8340
gaccctccc acattacggc ggagacggct aagcgttaggc tggccagggg atctcccccc 8400
tcctggcca gctcatcagc tagccagctg tctgcccctt ctttgaaggc aacatgact 8460
acccgtcatg actccccggc cgctgaccctc atcgaggcca accttctgtg gggcaggag 8520
atgggggggaa acatccccggc cgtggagtc gaaaataagg tagtaattttt ggactcttc 8580
gagccgtcc aacgggagga ggtatggggatcc ttccggggaa gatctcggt 8640
agttccagga aatttctcg akgatgccc atatggggc acggggatcc caacccctca 8700
ctgttagagt ctggaaagga cccggactac gtccctccaq tggtacacgg tggtccattg 8760
ccgcctgcca aggccccctc gataccacct ccacgggggg agaggacggg tgccctgtca 8820
gaatctaccg tgccttctgc ttggcggag ctgcgcacaa agaccttcgg cagctccgaa 8880
tcgtcgcccg tcgacacggg cacggcaacg gctctccctg accagccccctc cgacgacggc 8940
gacggggat ccgacgttga gtcgtactcc tccatggccc cccttgagggg ggagccgggg 9000
gatccccatc tcagggacgg gtcttggct accgttaagcg aggaggctag tgaggacgtc 9060
gtctgtctgtc cgatgtctca catatggaca ggccgcttga tcaacgcctg cgctcggt 9120
gaaaccaacg tccctgtcaa tgcgtgacg aactcttgc tccgtaccca caacttggtc 9180
tatgtacaa catctcgacg cgcacggctg cggcagaaga agtcaaccc ttggacactg 9240
caggctctgg acgaccacta cccggacgtg ctcaaggaga tgaaggcggaa ggcgtccaca 9300
gttaaggcta aacttctatc cgtggaggaa gcctgtaaac tgaccccccc acattcggycc 9360
agatctaaat ttggctatgg ggcaaaaggac gtccggaaacc tatccagcaa ggccgttaac 9420
cacatccgtt ccgtgtggaa ggacttgcgt gaaagacactg agacaccaat tgacacccacc 9480
atcatggcaaaaatggat ttctgcgtc caaccagaga agggggggccg caagccacgt 9540
cgccctatcg tattttccaga ttgggggtt cgtgtgtcg aaaaaatggc cctttacat 9600
gtggcttcca cccttccatca ggccgtatgg ggtcttcat acggatccca atactctct 9660
ggaacgggg tcgatgttcc ggtatggcc ttggaaacgca aaaaatggcc tatgggttc 9720
gcatatgaca cccgtgttt tgactcaacg gtcactgaga atgacatccg tttggaggg 9780
tcaatctacc aatgttgc tttggcccccc gaagccacac agggccataag tgctcataca 9840
gagccgtttt acatccccggg cccctgtact aattctaaag ggcagaactg cggctatccg 9900
cggtgcccg cgagccgtgt actgacgacc agtgcggta ataccctcac atgttatttg 9960
aaggccgtt ccgtctgtcg agtgcgaag ctccaggact gcacgatgct cgtatcgga 10020
gacgacccctt tcgttatctg tggaaacggcg gggacccaaag aggacgaggc gagectacgg 10080
gccttcacgg aggtatgtac tagatactct gccccccctg gggacccggc caaaaccggaa 10140
tacgacttgg agttgataac atcatgtctcc tccaatgtgt cagtcgcgca cgatgcacat 10200
ggcaaaaggg tttactatct cccctgtgac cccacccacc cccttgcggc ggtgtcggt 10260
gagacactgac gacacactcc agtcaattcc tggcttagca acatcatcat gtatgcggcc 10320
accttgggg caaggatgat cctgtactt cattttctt ccattcttctt atgtcaggaa 10380
caacttggaaa aagccctaga ttgtcagatc tacggggctt gttactccat tgagccact 10440
gacctaccc agatattca acgactccat ggccttagcg cattttactt ccatagttac 10500
tctccagggt agatcaatag gttggcttca tgcctcaggaa aacttggggt accgccccctt 10560
cgagtctgg aacatcgccc cagaagtgtc cggtcttaggc tactgtccca gggggggagg 10620
gctgcccattt gtggcaagta cctcttcaac tggcagtaa ggaccaagct caaactcact 10680
ccaaatcccg ctgcgtccca ttggattt tccagctgg tgggtctgg ttacagccgg 10740
ggagacat atccacagctt gtctcggtcc cgaccccggt ggttcatgtg tgccctact 10800
ctactttctg taggggttagg catctatcta cttcccaacc gataacggg gagctaaaca 10860
ctccaggccca ataggccatc ctgtttttt ccctttttt tttctttttt tttctttttc tttcccttgg 10920
ttttttttt tttttttttt ccctttttt tttctttttt tttctttttc tttcccttgg 10980
tggctccatc ttagccctag tcaacggctag ctgtgaaagg tccgtgagcc gttgactgc 11040
agagagtgtct gatactggcc tctctgcaga tcaagt 11076

<210> 8
<211> 8001
<212> DNA
<213> Hepatitis C Virus

<400> 8
gcccggcccc gättgggggc gacactccac catacatcac tcccctgtga ggaactactg 60
tcttcacgca gaaägctct aäccätggcg ttagtatgag tgcgtgcag cctccäggäc 120
ccccctcccc gggägagcca taigtggctg cgaaaccggg gagtacaccg gaattgccc 180
gacgaccggg tccttcttg gatcaäcccg ctcaatgcct ggagatttg gggtgcccc 240
gcgagactgc tagccgagta gtgtgggtc gcgaaaggcc ttgtggtaäct gcctgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcaccg aatccta 360
ctcaaägaaä aaccaaacgt aacaccaacg ggccgcaccat gattgacaaa gatggatgc 420
acgcaggatc tccggccgt tgggtggaga ggctattcg ctatgactgg gcacaacaga 480
caatcggtgt tcctgtatgc ggctgttcc ggctgtcagc ggaggggcgc cgggttctt 540
ttgtcaägac cgacctgtcc ggtgcctgta atgaactgca ggacgaggca ggcgggctat 600
cgtggcttgc cacgacgggc gttcttgcg cagctgtct ctagcttgc actgaaggcg 660
gaaggggactg gctgttattg ggcgäagtgc cggggcagga tctctgtca ttcacccctt 720
tcctgcggaa gaaagtatcc atcatggctg atgcaatgcg gggctgtcat acgcttgc 780
cggttacttg cccattcgac caccaägca aacatcgcat cgacgagca cgtactcgga 840
tggaaäggccg tcttgcgtat caggatgatc tggacgaa gcatcagggg ctcgcgc 900
ccgaaactttt cggccaggctc aaggcgccca tgcccacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcgttgcg äatatcatgg tggaaaatgg ccgttcttctt gattcatcg 1020
actgtggccg gctgggtgtg gggtggccgt atcaggatc acgcttggat acccgatcg 1080
ttgtctgaaga gcttggggc gaatgggctg accgttctt cgtgttttac ggtatcgccg 1140
ctcccgattt gcaacgcatac gccttctatc gccttcttgc ctagtttcc ccccccctt 1200
cagaccacaa cggtttccct ctacgcggat caattccgc cctctccctc tatgttattt 1260
aacgttactg gcccggccg ctggaaataa ggccgggtgtg ctttgtcttac tccaccatät tggctgtctt 1320
tggcaatgtt agggccggaa aacccgtggcc acggatctt ctttttttttac gttcttctt 1380
acggatctt ctggggctt tccccccttc gccaaäggaa tggacggatctt gttgaatgtc 1440
gtgaaaaag cagtcttgcgta gaaatggatc tggatggatc ctttttttttcc 1500
tgcaggcägc ggaacccccc acctggcgac gggtgtctt cggccaaaaa gccacgtiga 1560
taagatacac ctgcaaaaggc ggcaäcaaccc cagtgcacg ttgtgatgtt gatagtgt 1620
gaaagatgtca aatggctctc ctcaacgcgtt ttcacaaagg ggctgaagga tgcccagaäg 1680
gtacccatt gtatggatc tgatctgggg cctcggtca catgttttac atgtgttttac 1740
tcgagggtta aaaacgtctt gggcccccggaa accacggggc cgtgttttcc ctttggaaaa 1800
caccataataa ccatggcgcc tattacggcc tactcccaac agacgcgagg cttacttggc 1860
tgcatcatca ctagecctac agggccggac aggäaccagg tcgaggggga ggtccaatgt 1920
gttcäaccgg caacacatc ttccctggcg acctgcgtca atgcgtgtt gttgactgtc 1980
tatcatgggt cggctcaä gacccttgcg gggccaaagg gccaatcac ccaaatgtac 2040
accaatgtgg accaggacat cttccgttgcg caagcccccc cggggcgccg ttcttgcaca 2100
ccatgcacat gggcagctc ggaccttac ttgttcacga ggcacgcggc tgcattccg 2160
gtgcggccggc gggggcagacag caggggggacg ctactctccc ccaggcccgt ctcttactt 2220
aagggtctt cgggggttcc actgctctgc ccctcgccgc acgtgttggg catcttccg 2280
gtgcggctgt gcaacccgggg ggttgcgaag ggggtggact ttgtacccgtt ctagtctatg 2340
gaaäccacta tgcgggttccc ggtcttacgc gacaactgtt ccctccggc cgtaccgcag 2400
acattccatc tggccatcatc acacgcccctt actgttgcgtt gcaagacgcac taagggtgc 2460
gtgcgtatgc cggcccaagg gtataagggtt cttgtcttgc accgttccgtt cggccgcacc 2520
ctagtttgc gggcgatatac gtcttacgc catgttatc acccttacatc cagaatcägg 2580
gtaaggacca tcaccacggg tggcccccattt acgtacttcca cctatggca gtttcttgc 2640
gacgggtgggtt gctctggggg cgcctatgac atcataatat gtgtatgttgc ccactcaact 2700
gactcgacca ctatcttggg catcgccaca gtcctggacc aaggggagac ggctggaggcg 2760
cgactcgctg tgctcgccac cgtctacgcctt ccgggatcgg tcaccgtgcc acatccaaac 2820
atcgaggagg tggctctgtc cagcacttggc gaaatccctt tttatggcaaa agccatcccc 2880
atcgagacca tcaagggggg gaggcacccat attttctgc attccaagaa gaaatgtgt 2940
gagctcgccg cgaagctgtc cggccttgcgatc tcaatgttgc tagatcatcc cggggccctt 3000
gtatgtatccg tcataccaac tagcgaggac gtcattgtcg tggacacggc cgtctataatg 3060
acgggcttta cccgggtactt ctagtgcgtt atcgacttgc atacatgtgtt caccatcgac 3120
gtcgacttca gcctggacc gaccttccatc atttgagacca ctagccgtgcc acaagacgcg 3180
gtgtcacgtt cgcaggccg aggccaggact ggttagggca ggatgggcatttacaggat 3240
gtgactccag gagaacggcc ctcggccatc ttgcatttcc cgggttctgt ctagtgc 3300
gacgggggtt gtcgttgcgta ctagtgcgtt cccggccgaga cctcagtttgc gttggggct 3360
tacctaaaca caccagggtt gcccgttgc caggaccatc tggatgttgc gggagacgcgtc 3420
tttacaggcc tcaccacatc agacgcccat ttcttgcaccatc agactaagca ggcaggagac 3480
aactccctt acctgttgcg ataccaggatc acgggtgtcg ccagggtca ggttccacat 3540
ccatcggttcc accaaatgttgc gaaatgttcc atacgcgttcc accgttgcgtt gacccgttcc 3600

acgccccctgc ttgtataggct gggagccgtt caaaacgagg ttactaccac acaccccccata 3660
accaaataca tcatggcgtatgatgtccatgc catgtcggt gacctggagg tcgtcacgag cacctgggtg 3720
ctggtaaggcg gagtccttagc agctctggcc gctatttgcc tgacaacagg cagctggtc 3780
atttgtggca ggatcatctt gtccggaaag cggccatca ttcccgacag ggaagtcttt 3840
taccggagaatgt tcgtatgagat ggaagagtgc gcctcacacc caaacagaag gcaatcggt tgctcaaaac 3900
atgcagctcg cogaacaatt caagcgagg ctgctgtctcc gctggggaa tccaagtgcc tgctcaaaac 3960
gcgaagata tggaaattt catcagcggg atacaattt atcactatgt gcattcacag cttctatcatc 3990
cctggcaacc cggcgatagc accacccaaatc ataccctctt gtttacatc ctggggggat ttcctgtaggc gcccaccaag 4020
cctcccgatcg ctgcttctgc ataggctt ggaagggtgct tggttatatt ttggcaggtt ggaccatcga agccttctgg 4080
catcagcggg atacaattt atcactatgt gcattcacag cttctatcatc cagcccgctc 4140
cctggcggcc ctgtcgcc acgtggggccctt ttcctgtaggc gcccaccaatc ggggtggccgc ccaacttgc 4200
ttcgttcgc gggtaaccc gcaatcggt tggttatatt ttggcaggtt tggtggcggcc 4260
gcacgtgtca ctcagatctt ctgtcgcc acgtgggtt ggggtggccgc 4320
cagtggatca acggaggatcg gattggatatt gcaacgtgtt ggggtggccgc 4380
cgattggccgg ggttccctt gacggcatca tggaaaccac ggggtggccgc 4440
gggttccatga ggtatcggtt ggggtggccgc 4500
attaacgcgt acaccacggg ctgtggccgg tggctgtgt gactgatcc aagacctggc 4560
gtgacggccg tgaccactgatcgatctt ggggtggccgc 4620
ttcagacaag tggatggggatcgatctt ggggtggccgc 4680
cgggaggagg ttcacatctt tgccatgtt ggggtggccgc 4740
tgccggcccg aaccggatgt acggcgaga cggctaagcg 4800
tcagctagcc agctgtctgc cccgacgctg acctcatcg 4860
acccgcgtgg agtcagaaaa ggggtggccgc 4920
gaggaggatcgatctt ggggtggccgc 4980
cctcgacgca tgcccatatcgatctt ggggtggccgc 5040
aaggaccggg actacgtccc cctccgatatacctt ggggtggccgc 5100
cctccgatatacctt ggggtggccgc 5160
tctgccttgg cggagctcgatctt ggggtggccgc 5220
agcagtgtctc acttccatgtt ggggtggccgc 5280
taggctgccc acggatctc cttccacatcgatctt ggggtggccgc 5340
gccttcctt ggggtggccgc 5400
ggccaaccc tctggccggc 5460
taaggtatgtt ggggtggccgc 5520
atccgttccg ggggtggccgc 5580
ggcaccccg ggggtggccgc 5640
tccagtgtt ggggtggccgc 5700
gaggaaaggac ggggtggccgc 5760
cacaagagacc ttcggccatgtt ggggtggccgc 5820
tccctgaccatcgatctt ggggtggccgc 5880
gccccccctt ggggtggccgc 5940
aaggcgaggatcgatctt ggggtggccgc 6000
cctgatcatcgatctt ggggtggccgc 6060
tttgctccgtt ggggtggccgc 6120
gaagaaaggatcgatctt ggggtggccgc 6180
ggagatgatcgatctt ggggtggccgc 6240
taagctgatcgatctt ggggtggccgc 6300
gaacatatcc agcaaggccg ttaaccatcgatctt ggggtggccgc 6360
caactgagaca ccaatttgacccatcgatctt ggggtggccgc 6420
agagaaggatcgatctt ggggtggccgc 6480
gtgcgagaaaa atggccctt acgatgtgtt cttccacatcgatctt ggggtggccgc 6540
tccatcgatctt ggggtggccgc 6600
tttgcgatcgatctt ggggtggccgc 6660
tgcgatcgatctt ggggtggccgc 6720
cagacaggcc ataaggtcgatctt ggggtggccgc 6780
ggggggccccc tggatcgatctt ggggtggccgc 6840
gggtgactcgatctt ggggtggccgc 6900
tgtcgagatcgatctt ggggtggccgc 6960
atctgtggaaatcgatctt ggggtggccgc 7020
atgactatcgatctt ggggtggccgc 7080
ataacatcatcgatctt ggggtggccgc 7140
tatctccatcgatctt ggggtggccgc 7200
actccatcgatctt ggggtggccgc 7260
atgatctt ggggtggccgc 7320
cttagattgtc agatcgatctt ggggtggccgc 7380
attcaacatcgatctt ggggtggccgc 7440
aatagggtggatcgatctt ggggtggccgc 7500
cggggccagaa ggttcccgccgatctt ggggtggccgc 7560
aagtacatcgatctt ggggtggccgc 7620

tcccagtgg atttatccag ctgggtcggt gctggttaca gcgggggaga catatatcac 7680
agcctgtctc gtggcccgacc cccgctgggttc atgtggtgcc tactcctact ttctgttaggg 7740
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
ccatccgtt tttttccctt ttttttttcc tttttttttt tttttttttt 7860
ttttctccct ttttttccct ttttttttcc ttttttttcc ttgggtggct ccacatcttagc 7920
ccttagtcacg gctägctgtg aäaggccgt gagccgttg actgcagäga gtgctgatac 7980
tggccctctc gcagatcaag t 8001

<210> 9
<211> 11076
<212> DNA
<213> Hepatitis C Virus

<400> 9
gcacggcccc gattggggc gacactccac catagatcac tcccctgtga ggaactaactg 60
ttttcacgca gaaacgtct agccatggcg ttagtatgg tgctgtcag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaagggt gagtacaccg gaattgccag 180
gacgaccggg tccttcttg gatcaacccg ctcaatgcct ggagatttg gggtgcccc 240
gctgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgtact gcctgataagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaagaaa aaccaaacgt aacaccaacg ggccgcctcat gatggacaaa gatggatgc 420
acgcagggtc tcggccgtc tgggtggaga ggcttctcgatctactgg gcacaacaga 480
caatcggtcg ctctgtgtcc gccgtgttcc ggctgtcagc gcacggggcgc ccggttctt 540
ttgtcaagac cgacctgtcc ggtgcctgtaa atgaactgcg ggacgaggca ggcggctat 600
cgtggctggc cacgacgggc gttcctgtcg cagctgtgtc cgacgttgc actgaagcgg 660
gaaggagactg gctgttattt ggcgaagtgc cggggcaggaa tctcctgtca ttcaccc 720
ctcctgcgcg gaaagtatcc atcatggctg atgcaatgcg gcccgtcgt acgcttgatc 780
cggtactctg cccattcgat caccaagoga aacatcgat cgacgcggca cgtactcgga 840
tggaaaggccg tcttgcgtc caggatgtc tggaccaaga gcatcgaggc ctcggccag 900
ccgaaactt cgcggcgtc aaggcgcgcg tgcccacggc cggggatctc gtcgtgaccc 960
atggcgatgc ctgttgcgtc aatatcatgg tggaaaatgg cccgtttctt ggattcatcg 1020
actgtggccg gctgggtgtg cggggacccgt atcaggacat acgcttgcgt acccgatcg 1080
ttgtgaaga gcttggccg gaatgggctg accgttccct cgtgttttac ggtatcgccg 1140
ctcccgttcc gcaacgcata gccttctatc gccttcttgc cagtttctc tgatttaaa 1200
cagaccacaa cgggttccctt ctatcggttcaatccgc cctctccctc cccccccct 1260
aacgttactg gccgaagccg ctggaaataa ggcgggtgtg cttttgtcta tatgttattt 1320
tccacccat tggccgtttt tggcaatgtg aggccggga aacatggccc tgcgttctt 1380
acgacccatc ctatgggttctc tttccctctc gccaaggaa tgaacgggtct gttgaatgtc 1440
gtgaaaaag cagttctctt gaaatgttct tgaagacaaa caacgttgc tgcaggcggc 1500
ggaaacccccc acctggcgcac aggtgttctt gggggcggaaa gccacgtgt 1560
taagatacac ctgcaaaaggc ggcacaaccc cagtgcacgg ttgtgagttt gatagttgt 1620
gaaagagtca aatggcttc ctcaacgcgtt tcaacaagg ggctgaagga tggccagaag 1680
gtacccatt gtatggatc tggatctggg cctcgggtca catgttttac atgtgtttag 1740
tcgagggtttaaaaacgtctt gggcccccga accacggggc cgtgttttc ctttggaaaa 1800
cagcataataa ccatgggcac gaaatctaaa cctcaaaaggaa aaacccaaacg taacaccaac 1860
cgcccccac acgacccatc gtttggggc ggtgggtca ctcgtgggtt 1920
ttggccgcga gggggcccg gttgggtgtc cgcgcgacta ggaagactt cgaacgggtcg 1980
caacccctgtt gaaaggcaca acctatcccc aaggctcgcc agccggaggc tagggccttg 2040
gctcagcccg ggtacccctt gcccctctat gcaatggggc gcttgggtt ggcaggatgg 2100
ctcctgtcactt cccgtggc tggccctagt tggggccca cggacccccc gctgtgggtcg 2160
cgcaatttgg gtaaggcatcgatccatccctc acgtgcggct tcggcgtatc catgggttac 2220
attcccgctcg tggccgcggc cctagggggc gctgcaacggg ccctggcga tggcgtccgg 2280
gttctggagg acggcgttca ctatgcacca gggaaatctgc cccgttgc ttttctatc 2340
ttccttttttgg ctgttgcgtc ctgttgcgtt atcccgatctt cccgttgc ttttgc ttttgc 2400
gtatccggatgtatccatgttgcacgac gtcgttgc ttttgc ttttgc ttttgc ttttgc 2460
ggggacatgttgc ttcgttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 2520
cgctgttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 2580
acgatacgtac gccatgtcgatccatgttgc ttttgc ttttgc ttttgc ttttgc ttttgc 2640
gttggggatc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 2700
cgccacggaa ctttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 2760
cgatggctt gggatatgtt gatgttgc ttttgc ttttgc ttttgc ttttgc ttttgc 2820
tttactccggatccatgttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 2880
ggggcccttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 2940
ctctttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3000
ctcggttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3060
aacggcagctt gggatatgtt gatgttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3120
tttactccggatccatgttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3180
ggccagctgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3240
tcacacatgtt gggatatgtt gatgttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3300
gttaccggcg ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3360
ggggacggcc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3420
ctgttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3480
agcactgggtt ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3540
aaaaccttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc ttttgc 3600

tgggttcgg ggccttgggtt gacacccaga tgcttgggcc actaccata caggctttg 3660
 cactaccctt gcactgtcaa cttaaccatc ttcaaggtta ggatgtacgt ggggggagtg 3720
 gaacacaggc tcgaagccgc atgcaattgg actcgaggag agcgttgtaa cctggaggac 3780
 aggacaaatc gagacttag cccgtctgtc ctgtctacaa cggagtggca ggtattgccc 3840
 tgttccctca ccacccatc ggcttgcgtc actgggttga tccatctcca tcagaacgtc 3900
 gtggacgtac aatactgtt cggataggg tggcggttgc ttccttgc aatcaaattgg 3960
 gagtatgtcc tggttgcctt ctttcttc gggacggcgc ggttgcgtc ctgtttgttgc 4020
 atgatgtgc tgatagctca agctgaggcc gccctagaga acctgggtgtt ctcacacgtc 4080
 gcacccgtgg cggggcgca tggatttcc tccttctcg tggttcttgc tgctgcctgg 4140
 tacatcaagg gcaggctggt ccctggggcg gcatatgccc tctacggcgt atggccgcta 4200
 ctccgtctcc tgctggcgtt accaccacga gcatacgcca tgaacccggga gatggcagca 4260
 tgggtggag ggcgggtttt cgttagtctg atactcttgc ctttgcacc gcactataag 4320
 ctgttccctcg cttaggtctat atgggttgtt caatattttt tcaccaggcc cgaggcacac 4380
 ttgcaagttt ggatcccccc ctcacacgtt cggggggggcc gcatatgttgc datccctctc 4440
 acgtgcgcga tccacccagaa gtaatctt accatccaca aaatcttgc cgcctactc 4500
 ggtccactca tgggtctca ggctggata accaaaatgc cgtacttcgt ggcgcacac 4560
 gggcttcatc gtgcgtcat gctgggtgcgg aaggttgcgtt ggggtcatta tgtccaaatg 4620
 gctctatcgtt agttggccgc actgacaggt acgtacgtt atgaccatct caccctactg 4680
 cgggacttggg cccacccggc cttacggac tttgcgttgc cagttgagcc cgtcgttcc 4740
 tctgtatgg agaccaaagg tatcacctgg ggggcagaca ccgcggcgtt tggggacatc 4800
 atcttggcc tggcggttcc cggccgcagg gggagggaga tacatctggg accggcagac 4860
 agccttgaagg gcaagggttgc ggcacttccgc ggccttattt cggcctactc ccaacagac 4920
 cggggcttac ttgggtcat catactatc ctcacaggcc gggacaggaa ccaacatc 4980
 ggggggggtcc aagtggtctc caccgcacaca caatcttcc tggcgacatc cgtcaatggc 5040
 gtgtgttggta ctgtctatca tggttgcggcc tcaaagaccc ttggcgccccc aaaggccca 5100
 atcaccctaaa tggatcacca tgggaccac gacctcgcc gttggcaagg gccccccggg 5160
 ggcgcgttcc tggatccatc cacccgtggc agtcggacc tttacttggg caccggcat 5220
 ggcgcgtca ttccgggtcg cccggggggc gacagcaggg ggagctactt ctcccccaagg 5280
 cccgtctctt acttggaaagg ctcttggggc ggtccacttc tctggcccttc ggggcacatc 5340
 gtgggcattt ttcgggttgc ctgtgtcacc cgggggtttt cggaggccgtt ggactttgtt 5400
 cccgtcgagt ttagtggaaac cactatgcgg tccccgtct tcacggacaa ctcgtccct 5460
 cccggcgatc cgcagacattt ccagggtggc catctacacg cccctactgg tagcggcaag 5520
 agcactaagg tggccgttc gtagcgtaccc caagggtata aggtgttgc cctgaaccc 5580
 tccgtcgccg ccaacccatgg ttccggggcg tatatgtcta aggacatgg ttcgacatc 5640
 aacatcagaa tggggtaag gaccatcacc acgggtcccccc ccatcactgaa ctccacat 5700
 ggcacggttt ttggccacgg tgggttgcctt gggggccctt atgacatcat aatatgtat 5760
 gagttccactt caactgtactt gaccatctt ctggccatcg gcacagtccctt ggaccaagcg 5820
 gagacggctg ggcgtactt cttccgttgc gccaccgtt cccctccggg atcggtcaacc 5880
 gtggccacatc caaacatcgaa ggagggttgc ctgtccacga ctggagaaat ccccttttat 5940
 ggcacggccca tccccatcgaa gaccatcaag ggggggaggc accttattttt ctggcattcc 6000
 aagaagaaat gtgtatggact cggccgtcaag ctgtccggcc tggacttcac tgctgttagca 6060
 tattaccggg gccttgcgtt atccgtcata ccaacttagt gacgttgcact cttccatata 6120
 acggacgttc taatgacggg ctttaccggt gacttcgtact cttccatata ctgtcaatata 6180
 tgggttcaccc agacatgtca ctccatgttgc gacccgttcc ttcaccatttgc gacgacgacc 6240
 gtggccacaaac acggcggttc acgttcgtccatcg cggccggggca ggacttggtag gggcaggatg 6300
 ggcatttaca ggtttgttgc tccaggagaa cggcccttccg gcatgttgc ttccctcggtt 6360
 ctgtgcgttgc gctatgtaccc ggggtgttgc tggtacggc tcaacccggc cgagacctca 6420
 gtttaggttgc gggcttaccctt aacacacca ggggttcccg ttcggccagg ccatctggag 6480
 ttcttgggaga ggcgttttac aggccttccacc cacatagacg cccattttt gtcccttgcgtt 6540
 aagcaggccag gagacaactt ccccttacatcg gtagcatacc aggttacgggt gtgcggccagg 6600
 gtcaggccatc caccatccatc tggggaccaa atgtggaaatgttgcatacg gctaaagctt 6660
 acgtgtccatcg ggcacccatcc cctgtctgtt aggttggggccatccatcg ggggttact 6720
 accacacaccc cccatcataacc atacatcatcg gcatgtatcg cggctgttcc ggggttgcgtc 6780
 acggacatcg ggggtgttgc agggccatcg ctggccatcg tggccgttcc ttcaccatcc 6840
 acaggccatcg tggtccatcg gggccatcg atcttgcgttcc gaaagccggc catccatccc 6900
 gacaggccatcg tcccttccatcg ggatgttgc gatgtggaaatgttgcatacg acaccccttcc 6960
 tacatcgatcc accggatgtca gtcgtccggaa caatttccatcg ggggttgcgttcc 7020
 caaacacccca ccaacccatcc accggatgtca gtcgtccggaa tggatccaa ggggttgcgttcc 7080
 atcgttgcgttcc tcccttccatcg gcatatgttgc aatttccatcg gggatccatcg atatttgcgttcc 7140
 ggcgttgcgttcc tcccttccatcg caacccatcc gatgttgcgttcc tcccttccatcg gggatccatcg 7200
 atcaccatcg cccatcgatcc accggatgtca gtcgtccggaa tggatccaa ggggttgcgttcc 7260
 gccggccatccatcg tcccttccatcg gcatatgttgc aatttccatcg gggatccatcg gggatccatcg 7320
 ggggttgcgttcc tcccttccatcg gcatatgttgc aatttccatcg gggatccatcg gggatccatcg 7380
 gcaaggccatcg cccatcgatcc tcccttccatcg gcatatgttgc aatttccatcg gggatccatcg 7440
 gaggacatcg tcccttccatcg gcatatgttgc aatttccatcg gggatccatcg gggatccatcg 7500
 gtgtgtccatcg gcatatgttgc tcccttccatcg gcatatgttgc aatttccatcg gggatccatcg 7560
 aaccggatcgatcc tcccttccatcg gcatatgttgc aatttccatcg gggatccatcg gggatccatcg 7620

gagagcgcacg ctgcagcacg ttgtcaactcg atccctctcta gtcttaccat cactcagctg 7680
 ctgaagaggc ttccaccagtg gatcaacgag gactgctcca cgccatgcctc cggctcggtg 7740
 cttaagagatg tttggatttg gatatgcacg gtgttactg atttcaagac ctggctccag 7800
 tccaagctcc tgccgcatt gctgggagtc cccttcttct catgtcaacg tgggtacaag 7860
 ggagtctggc ggggcacgg catcatcaa accacctgcc catgtggggc acagatcacc 7920
 ggacatgtga aaaacgggtt catgaggatc gtggggctta gacactgttag taacacgtgg 7980
 catggacat tccccattaa cgcgtacacc acggggccct gcacgcctcc cccggcgcca 8040
 aattatctca gggcgctgtg ggggggtggct gctgaggagt acgtggaggt tacgcgggtg 8100
 ggggatttcc actacgtgac gggcatgacc actgacgacg taatgtgccc gtgtcagggt 8160
 cccggcccccgg aattttcac agaagtggat ggggtgggt tgacacaggta cgctccacg 8220
 tgcaaaaaaaa ttctacggga ggaggtcaca ttccctggteg ggctcaatca atacctgggt 8280
 gggtcacagc tcccatgcga gccccaaacgg gatgtacgc tgctcaactc catgtctacc 8340
 gacccctccc acattacggc ggagacggct aagcttaggc tgccagggg atctcctccc 8400
 cccttggcca gctcatcagc tagccagotg tctgcgcctt ccttgaaggc aacatgact 8460
 acccgctatg actdcccggc cgctgaccc tc acggggcc acctcttgcg gggcaggag 8520
 atggggggga acatcaccgg cgtggagtc gaaaataagg tagtaattt ggactcttc 8580
 gagccgtcc aagcgagga ggtatgggg gaagtatccg tttccgggaa gatctcggt 8640
 aggtccagga aattccctcg agcgtacgc atatggcac gcccggattaa caaccctcca 8700
 ctgttagagt ccttggaaaggc cccggactac gtccctccag tggtagacacgg gtgtccattg 8760
 ccgcctgcca aggccctcc gataccactt tcacggagga agaggacggg tgcctgtca 8820
 gaatctaccg tttttctgc cttggcgag ctgcacccag agacccctgg cagctccgaa 8880
 tcgtcggccg tcgacagcggc cacggcaacg gcctctctg accagccctc cgacgacggc 8940
 gacgcgggat cgcacgtga gtctgtactcc tccatggccc cccttgagggg ggagccgggg 9000
 gatccgatc tcagcgacgg gtcttggctt accgtaagcg aggaggctag tgaggacgtc 9060
 gtctgtctc ctatgttccat cacatggac ggcgccttcg tcaacggcatg cgctgcggag 9120
 gaaacaaga tgccatcaa tgactgagc aacttcttc tccgttccacca caacttggtc 9180
 tatgtacaaa catctcgag cgcaaggctc cggcagaaga agtgcacctt tgacagactg 9240
 caggctctgg aacgaccacta cccggacgtg ctcaaggaga tgaaggcgaa ggcgtccaca 9300
 gttaaaggcta aacttctatc cgtggaggaa gcctgtaaac tgacccccc acattcgcc 9360
 agatctaaat ttggctatgg gcaaaaggac gtccgaaacc tatccagcaa ggcgttaac 9420
 cacatccgct cccgttgaa ggacttgcg gaagacactg agacccaaat tgacaccacc 9480
 atcatggcaa aaaatggat tttctgcgtc caaccagaga agggggggccg caagccagct 9540
 cgccttatcg tattccaga tttttttttt cgtgtgtcg agaaaatggc cctttacat 9600
 gtggcttcca ccctccctca gggcgatg ggccttcat accgattcca atactcttc 9660
 ggacagcgcc tcgagttctt ggtatggc tggaaagcg agaaaatggcc tatggcttc 9720
 gcatatgaca cccgctgttt tgcactcaacg gtcaactgaga atgacatccg tggtagaggag 9780
 tcaatctacc aatgttgcg tttggccccc gaagccagac agggccataag gtcgctcaca 9840
 gagccggctt acatcggggg cccctgtact aattctaaag ggacaaactg cggctatcgc 9900
 cgggtcccgcc cgagcggtgt actgacgacc agtgcggta ataccctcact atgttactt 9960
 aaggcccgct cggctgtcg agtgcgaaag ctccaggact gcaacgtgtc cgatgcgg 10020
 gacgcacccgg tctgtatcg tggaaagcg gggacccaaag aggacggagc gagectacgg 10080
 gccttcacgg aggtatcg tagatacttcttcccttggggcccaaaaccggaa 10140
 tacgacttgg agttgataac atcatgttccatccatgtgt cagtcgcga cgatgcacat 10200
 ggcaaaagggg tgcactatctt caccgtgac cccacccaccc cccttgcg 10260
 gagacagacta gacacactcc agtcaattcc tggctaggca acatcatcat gtatgcgccc 10320
 accttgggg caaggatgtat cctgtatgact cattttcttccatccatgtgttccat agtgcggaa 10380
 caacttggaa aagcccttaga ttgtcagatc tacggggctt gttactccat tggccactt 10440
 gacctaccc agatcattca acgactccat ggccttagcg cattttactt ccatagttac 10500
 tctccagggtt agatcaatag ggtggcttca tggcttcggaa aacttgggtt accggccctt 10560
 cgagtctggc gacatcgccg cagaatgttc cggcgatggc tactgtccca gggggggagg 10620
 gctgcccactt gtggcaagta ccttcttccat tggcgatggaa ggaccaagct caaactact 10680
 ccaatcccg ctgcgtccca gttggatttttccatgtgttccatgtgttccat tttttttttt 10740
 ggagacatataatc acacagccctt gtcgtgcgcg acggccctt ggttcatgtg gtgcctactc 10800
 ctactttctg tagggtagg catctatcta ctcccaacc gatgaacggg gagctaaaca 10860
 ctccaggccca ataggccatc ctgtttttttt ccctttttt tttttttttt tttttttttt 10920
 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 10980
 tggctccatc ttggcccttag tcaacggctag ctgtgaaagg tccgtgagcc gcttgcgtc 11040
 agagactgtgtc gatactggcc tctctgcaga tcaagt 11076

<210> 10
<211> 8001
<212> DNA
<213> Hepatitis C Virus

<400> 10
gcccggcccc gattgggggc gacactccac catagatcac tccccgtgaa ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccccctccc gggagagcca tagtggctcg cggAACCGGT gaggacaccg gaattgccag 180
gacgaccggg tccttcttg gatcaacccg ctcaatgcct ggagatttgg gctgtccccc 240
gcgagactgc tagccgaga gtgttgggtc gggaaaggcc ttgttgtact gcctgtatagg 300
gtgttgcga gtgccccggg aggtctgcac gaccgtgcac catgagcacg aatccctaaac 360
ctcaaagaaa aaccAAACGT aacaccaacg ggcgcgcct gatggattaca gatggattgc 420
acgcagggtc tccggccgt tgggtggaga ggcttctgg ctatgactgg gcacaacaga 480
caatcgctg ctctgtatcc gccgtgttcc ggctgtcagc gcaggggcgc cgggttttt 540
ttgtcaagac cgacctgtcc ggtgcctgaa atgaaactgca ggacgaggca ggcggctat 600
cgtggctggc cacgaCGGJC gttccttgcg cagctgtct cgacgttgc actgaagcgg 660
gaaggggactg gctgttattg ggcgaagtgc cggggcagga tccctgtca ttcacccctg 720
ctccctgcga gaaatgtcc atcatggct atgaaatgcg gcccgtcat acgcttgatc 780
cggtactcg cccatcgac tccaaagega aacatcgcat cgagcgagca cgtactcgga 840
tggaaaggccg tcttgcgtat caggatgatc tggacaaaga gcatcggggg ctcgcggccag 900
ccgaactgtt cggcaggctc aaggcgcga tgccgcggg cggatgttc gtcgtgaccc 960
atggcgatgc ctgttgcgt aatatcatgg tggaaaatgg ccgtttttct ggtttcatcg 1020
actgtggccg gtcgttgcgt gcccggcgt atcaggacat agcgttgcgt acccggtata 1080
ttgtcaaga gcttggcgcc gaatgggctg accgttctt cgtgttttac ggtatcgccg 1140
ctcccgattc gcagcgcatac gccttctatc gccttcttgc ctaggttctt ctaggtttaaa 1200
cagaccacaa cgggttccct ttagccggat caatccgcct cctctccctc cccccccct 1260
aacgttactg gccgaaggcc cttggaaataa ggcgggtgtg cttttgtcta tatgttattt 1320
tccacccat tggcccttgc tggcaatgtg aggggccggaa aacctggccc tggcccttgc 1380
acgagcatc cttaggggtct tttccctctc gccaääggaa tcaaggcttct gttgaatgtc 1440
gtgaaggaaag cagtttcttgc gaaagcttct tgaagacaaa caacgtctgt akgcaccctt 1500
tgcaggcgc ggaacccccc acctggcgac aggtgcctct gcccggccaaa gccacgtgt 1560
taagatáac ctgcaaaaggc ggcacaaccc cägtggccacg ttgtgagttg gatagtgt 1620
gaaagagtca aatggctctc ctcaagcgta tcaacaaagg ggctgaagga tggccagaag 1680
gtatcccatt gtatggatc ttagtctgggg cctcggtgca catgttttac atgtttag 1740
tcgaggtaa aaaaatctca gcccccccgaa accagggggaa cgtgggtttt ctttgaaaaa 1800
cagcatataa ccatggcgc tattacggcc tactcccaac agacgcgaagg cctacttggc 1860
tgcataatca ctggctcac aggccgggac aggaacccagg tggagggggaa ggtccaaatg 1920
gtctccaccc caacacaatc ttcctggcg acctgcgtca atggcgatgt ttggactgtc 1980
tatcatgggtt cgggtcaaa gacccttgc ggcggccaaagg gcccaatcac ccaaatgtac 2040
accaatgtgg accaggacat cgtcggtctg caagcgcccc cggggggcgcg ttcccttgc 2100
ccatgcacct gggcagctc gaccccttac ttggtcacga ggcacatggca tgcattccg 2160
gtgcgcggc gggcgcacag caggggggacg ctactctcc ccaaggccgt ctcctactt 2220
aagggtctt cgggggttgc actgtctgtc ccctcgggc acgtgttggg catcttccg 2280
gtgcgtgt gCACCCGGGG ggttgcgaag gggtgtgact ttgttccctgt cgagtctatg 2340
ggäaccacta tgggttccct ggttcttgc gacaactgtt cccctccggc cgtaccggcag 2400
acattccagg tggcccatct acacccccctt actggtagcg gcaagacggc taagggtccg 2460
gtgcgtatg caggccaagg gtataagggtt ttgttcttgc acccggtccgt cgccggccacc 2520
ctagggttctg ggggtatata tgcataaggca catggatgtc accctaataat cagaatcgcc 2580
gtaaaggacca tcaaccacggg tgcccccatac acgtactcca cctatggca gtttcttgc 2640
gacgggtgtt gtcgttgggg cgcctatgac atcataatat gtgtatgtt ccactcaact 2700
gactcgacca ctatcttggg catcgacca gtcgttgcacc aagcgagac ggctggagcg 2760
cgactcgatc tgcgtccac cgctacgcct cgggatcg tcaccgtgc acatccaaac 2820
atcgaggagg tggctctgtc cagcgttgc gaaatccccctt tttatggca agccatcccc 2880
atcgagacca tcaagggggg gaggccatctt attttctgc attcaagaa gaaatgtgt 2940
gagctcgccg cgaagctgtc cggccttgcg ctcataatgtc tagatattt cggggccct 3000
gtgtatccg tcaatccaaatc tagcgagac gtcattgtcg tagacacggc cgtctaatg 3060
acgggttta cggcgactt cgtactcgatc atcgactgca atacatgtgt caccctacaca 3120
gtgcacttca gcctggaccg gacccatccac attgagacca cgaccgtgc acaagacgcg 3180
gtgtcacttca cgcaggccgg aggcaggact ggttagggca ggatggcat ttacaggtt 3240
gtgaéccag gagaacggcc ctccggcatg ttccatttctt cgggttctgt cgagtgtat 3300
gacgcgggtt gtgttggta cggactcaccg cccggccaga cctcgtttag gttgcgggtt 3360
tacctaaacaa caccagggtt gcccgttgc caggaccatc tggatgttgc ggagagcgtc 3420
tttacaggcc tcaaccacat agacccccat ttcttgcctt agactaaagca ggcaggagac 3480
aacttccctt acctgttgcg ataccaggct acgggttgcg ccagggtca ggctccacact 3540
ccatcggttcc accaaatgtt gaaatgttgc atacggctaa agcgttgcgtt gacccgttca 3600

acggccctgc tgtataggct gggagccgtt caaaacgagg ttactaccac acacccata 3660
 accaaataca tcatggcata catgtcggt gacctggagg tcgtcacgag cacctgggt 3720
 ctggtaggcg gagtcctagc agctctggcc gcgtattgcc tgacaaacagg cagcgtggc 3780
 attgtggca ggatcatctt gtccggaaag ccggccatca ttcccgacaig ggaagtccct 3840
 taccggagat tcgatggagat ggaagagtgc gcctcacacc tccttacat cgaacaggg 3900
 atgcactcg cccgaaacatt caaacagaag gcaatcggt tgctgcaaaag aecaccagg 3960
 caagcgagg ctgctgctcc cggtggaa tccaatggc ggaccatcg aeccttcgg 4020
 gcgaaacata tggatatt catcagcggg atacaatatt tagcaggctt gtccactcg 4080
 cctggcaacc ccggatagc atcaactgt gcattcacag cctctatc 4140
 accacccaaac ataccctctt gtttaacate ctggggggat ggtggccgc ccaacttgct 4200
 cctcciaagcg ctgcttctgc ttcgtatggc gcccgcacg ctggagcggc tggcagc 4260
 ataggccctt ggaaggctt tggatatt ttggcagggtt atggagcagg ggtggcaggc 4320
 ggcgtcgcc ccttaaaggat catgagcggc gagttggccct ccaccggagga cctggtaac 4380
 ctactccctg ctatcccttc ccctggccg ctagtcgtcg gggtcggtg cgcagcgata 4440
 ctgcgtcgcc acgtggccccc aggggggggg gctgtcagt ggatgaaacc gctgatagc 4500
 ttgcgttcgc gggtaacca cgtctcccc acgactatg tgccgtagag cgcgtcgca 4560
 gacgagtc ctcagatctt ctcgtatctt accatcactc agtgcgtgaa gaggttcac 4620
 cagtggatca acgaggatcg ctccacgc tgcgtccggc ctggctaaag agatgttgg 4680
 gattggatat gacgggtgtt gactgattt aagacctggc tccagtc 4740
 cgattggccgg gaggccctt ctctctatgt caacgtgggt acaagggagat ctggcggg 4800
 gacggcatca tgcacccac ctgcccattt ggagcacaaga tcaaccggaca tggaaaaaac 4860
 ggttccatgtt ggatgtggc gcttagggc tggatgttaca ctggcatgg aacattcccc 4920
 attaacgcgtt acaccacggg cccttcaccc cctcccccgg ccggaaattt ttctaggcg 4980
 ctgtggccgg taggtgttgc gggatgttgc ggggtggg 5040
 gtgacgggca tgaccactga caacgtaaag tggccgtgtc aggttccggc ccccgaaattt 5100
 ttcaacaaag tggatgggtt ggggttgcac aggtacgtc cagcgtgca acccctcta 5160
 cggggaggagg tcacatttccctt ggtcgccgtt aatcaatacc tggggggc acagctccca 5220
 tgcgagctt aaccggatgt agcagtgtc acttccatgc tcaccgaccc ctcccacatt 5280
 acggccggaga cggctaagcg taggtgtggc agggatctc ccccccctt ggccagctca 5340
 tcagctatcg agctgtctgc cccttcctt aaggcacaat gcaactaccg tcatgactcc 5400
 ccggacgtcg acctcatcg gccaacactt ctgtggccggc aggagatggg cgggaaacatc 5460
 accccgttgg agtcagaaaa taaggttagt attttggact ctggcgaccc gctccaaagcg 5520
 gaggaggatg agaagggagat atccgttccg gggagatcc tgccgaggto cggaaattt 5580
 cctcgagcga tgccatatg gcaacggccg gattacaacc ctccactgtt agagtcctgg 5640
 aaggacccgg actacgtccc tccagtggta cacgggtgtc cattggccgc tgccaaggcc 5700
 cctccgatatac cacccatcg gggaaagagg acgggttgc tgcgtatcc taccgtgtct 5760
 tctgccttgg cggactcgcc cacagaccc tccggcactt ccgaatcgatc ggccgtcgac 5820
 acggcacgg caacggctt tccgtaccag ccctccggc acggcgcacgc gggatccgac 5880
 gttgactcgat aactccatcat gccccccctt gggggggage cggggatcc cgatctc 5940
 gacgggttccctt ggttaccgtt aagcgaggag gctagtgggg acgtcgatcg ctgtcgat 6000
 tcctacacat ggacaggcgc cctgtatcc ccatcgctg cggagggaaac caagtcggcc 6060
 atcaatgcac tggacaactt tttgtccgtt caccacaact tggctatgc tacaacatct 6120
 cgcagcgcac acctgcccga gaagaaggcc acctttgaca gactcgagg tctggacgac 6180
 cactaccggg acgtgtccaa ggagatgaag gcaaggcgtt ccacagttt ggctaaactt 6240
 ctatccgtgg aggaaggctt taagctgtacg ccccccacatt cggccagatc taaatttggc 6300
 fatggggccaa aggcgtccg gaaacatcc agcaaggccg ttaaccacat ccgctccgtg 6360
 tggaaaggact tgcgtggaga cactgagaca ccaatttgcac ccaccatcat ggcaaaaaat 6420
 gaggttttctt ggttccaaacc agagaaggggg gggccaaaggc cagtcgatc tategtatcc 6480
 ccagattttgg ggggttgcgtgt gtcgagaaaa atggccctt acgtgtgtt ctccaccctt 6540
 cctcaggccg tgcgtggctt ttcataccgg tccaatactt ctcctggaca ggggtcgag 6600
 ttccctggta atgcctggaa agcgaagaaaa tggccctatgg gcttcgcata tgacaccggc 6660
 tggggactt caacgtcact tggaaatgtt atccgtgtt agggtcaat taccatgtt 6720
 tgcgtactgg ccccccggcagc cagacaggcc ataaggctgc tcaacagacg gctttacatc 6780
 gggggccccc tgcataatcc taaagggcag aactcgccgtt cggccgcgagc 6840
 ggtgtactga cggccatcg cgttaatacc ctcacatgtt accttggacgc cgctcgccggc 6900
 tgcgtcgactt cggacttccaa ggactcgacg atgcgttat gggagacga ctttgcgtt 6960
 atctgtgaaa ggcgggggac ccaagaggac gggccggcc tacggccctt caccggatc 7020
 atgacttagat actctggccc ccctggggac ccggccaaac cagaatacga tggagttt 7080
 ataacatcat gtcctccaa tgcgtatcc ggcacgtat catctggcaa aagggtgtac 7140
 tatctccatcc tgcacccac cccccccctt gggccggccgtt cgtggagac agcttagacac 7200
 actccactca atccgtggctt agcaacatc atctgtatcg cggccacccctt gtggccagg 7260
 atgatctgtt tgcataatcc ctctccatcc ttcgtatcc tggatgttgc cggccacccctt 7320
 ctagatgttgc agatcttccgg ggcgttgc tccattggc cactgtaccc acctcagatc 7380
 atcaacgac tccatggctt tagcgatcc tcaactccata gttactctcc aggtgagatc 7440
 aatagggtgg ctcatgttcc caggaaactt ggggtaccgc ctttgcgtt gggagacatc 7500
 cggcccgagaa ggttccggc taggtactt tcccgagggg ggaggccgtc ctttgcgtt 7560
 aagtactctt tcaactggc agtaaggacc aagtcacaaat cccggctcg 7620

tcccagggtgg atttatccag ctggttcggt gctgggttaca gcgggggaga catatatcac 7680
ägcctgtctc gtgcccgaacc ccgcgtggtc atgtggtgcc tactcctact ttctgttaggg 7740
gtaggcatct atctactccc caaccgtatga acggggagct aaacactcca ggccaatagg 7800
ccatccgtt tttttccctt ttttttttcc tttttttttt tttttttttt 7860
ttttctccctt tttttttctt ctttttttcc ttttttttcc tttggtggct ccatcttagc 7920
cctagtccacg gctagctgtg aääaggccgt gagccgcttg actgcagaga gtgctgatac 7980
tggccctcttgcagatcaag t 8001

<210> 11
<211> 11076
<212> DNA
<213> Hepatitis C Virus

<400> 11

gcccggcccc gáttggggc gacactccac cáttagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaágctct agccatggc ttatgtatgag tgctcgta cctccaggac 120
ccccctcccc gggagagcca tagtggctg cgaaaccgg gatgtacaccg gaattggca 180
gacgaccgg gtccttcttg gatcaaccg ctcaatgcct ggagatttg gctgtcccc 240
gctgagactgc tagccgaga gtgtgggtc gctgaaaggcc ttgtggtaact gcctgtatagg 300
gtgttcgca gtgccccggg aggtctcgta gaccgtgcac catgagcact aátctaaac 360
ctcaaagaaa aaccaaacgt aacaaccacg ggcgcgcccatt gattgaacaa gatggáttgc 420
acgcaggttc tccggccgt tgggtggaga ggcttattgg ctatgactgg gcaacaacaga 480
caatcgctg ctctgatgcc gcccgttcc gctgtcagc gcaaggggcgc cccgttctt 540
ttgtcaagac cgacctgtcc ggtgcctcta atgaaatgc ggcaggggcgc gctgggttat 600
cgtggctgca caccacggcc gttccttgcg cagctgtctc cgacgttgtc actgaagcgg 660
gaaggactg gctgttattg ggcgaatgtc cggggcagga tccctgtca ttcacccctt 720
ctccctggca gaaagtatcc atcatggctg atgcaatgc gcccgtcat acgcttgatc 780
cggtctactg cccatcgac caccacggca aacatcgat cggcggcagca cgtactcgg 840
tggaaaggc tcttgcgt caggatgtc tggacggaga gcatcaagg 900
ccgaaactgtt cgcggccgtc aaggcgcgc tggccggcgg cggaggatctc gtcgtgaccc 960
atggcgatgc ctgttgcgg aataatcátgg tggaaatgg ccgttcttcc ggatttcatcg 1020
actgtggccg gctgggtgtg gcccggccgt atcaggacat acggttgcg accegtata 1080
ttgtcaaga gcttggccgc gaatgggtc accgttctc cgtgttcttac ggtatcgcc 1140
ctcccgattc gcagcgatc gccttctata gccttcttgc caggttcttcc tgatgtttaaa 1200
cagaccacaa cgggttccct ctacggggat caatccgc cctctccctc cccccccct 1260
aacgttactg gccgaaggcc cttggataaa ggccgtgtg cgttgcata tatgttattt 1320
tccacccat tgcgttctt tggcaatgtg agggccggaa aacctggccc tgcgttctt 1380
acgaggattc ctgggggtc tttccctctc gccaaaggaa tgcgttgcgt gttgaatgtc 1440
gtgaaggaaag cagttcttgc ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggccgc ggaacccccc acctggcgc acgttgcctc gcccggccaa gccacgtgt 1560
taagatacac ctgcaaaaggc ggcacaaccc cagtgcacg tttgagttg gatgttgc 1620
gaaagactca aatggcttc tcaacgcgtt tcaacaagg ggctgttgcg tgcccagaag 1680
gtacccatt gtatggatc tgatctgggg cctcggtca catgttttac atgtgtttag 1740
tcgagggttaaaaacgtcta ggccccccca accacggggc cgtgttttc ctttggaaaa 1800
cäcgataataa ccatgggcac gaatctaaa cctcaaaagaa aaaccaaacg taacaccaac 1860
cggcccccac aggacgtcaa ttggccggg ggtgttgcg tgcgttgcgtt agtttacctg 1920
ttggccgcga ggccccccca gttgggtgtg cggcgtactt ggaagacttgc gtagcggtcg 1980
caacctctgtt gaaaggcaca acctatcccc aaggctcgcc agcccgagg tagggcttgg 2040
gctcagcccg ggttcccttgc gcccctctat ggcaatgtt gcttgggtt ggcaggatgg 2100
ctccctgttcc cccgtggctc tggccctatg tggggccca cggacccccc gctgttgcgt 2160
cgcaatttgg tgaaggatc ctataccctc acgttgcgtt tgcgttgcgtt catgggttac 2220
atccctctgtc tggccccc cctagggggc gcttgcgtt ccttggccca tggcgtccgg 2280
gttctggagg acggcgttca ctatgcaca gggaaatctgc ccttgggttctc cttttctatc 2340
ttccctttgg ttttgcgtc ctgttgcgtt atcccttgcgtt ccttgcgttgcgtt agtgcgcääac 2400
gtatccggag ttttgcgttgcgtt ccttgcgttgcgtt ccttgcgttgcgtt cttttctatc 2460
gcccggatgtca tcatgcatac ccccccgggtc gtcccttgc ttccggagaa caactctcc 2520
cgctgttggg tagtgcgttca tcccaacgtc gcccggccaggaa acgttgcgtt ccccaactacg 2580
acgataacac gccatgttca ttgttgcgtt gggggccgtt ctcttgcgtt cgttatgtac 2640
gtggggatc tctggggatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 2700
cggttccatgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 2760
cgatgttgcgtt gggatgtatgat gatgttgcgtt tcacccatgtt cttttctatc ttttgcgttgcgtt cttttctatc 2820
ttactccggat tcccaacgtc ttgttgcgttgcgtt atgggtgggg gggcccttgc gggggatgttgcgtt cttttctatc 2880
gcccccccttgcgttccatgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 2940
ctctttggatc gcttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3000
ctccggatgttca ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3060
aacggcgttca ggcacatcaa caggactgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3120
ttcccttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3180
gcccggccatgtca gcccggccatgtca cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3240
tcacacatgttca gggacccatgtca ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3300
gttcccttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3360
ggggacccatgtca gggacccatgtca ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3420
ctgttgcgttca acaacacgtc gcccggccatgtca ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3480
agcacttgggt tcaacaaacgtc gttcccttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3540
aaaacccatgtca cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc ttttgcgttgcgtt cttttctatc 3600

Vainberg, Irina

Fr: Fehlner, Paul
Sent: Tuesday, September 30, 2003 8:06 AM
To: 'Michael Sasaki, Esq. (msasaki@anadyspharma.com)'
Cc: Vainberg, Irina
Subject: Hepatitis C Virus Constructs ... (PCT/US01/46350; our file 00342/200h395-WO0)

Michael:

In the absence of instructions to the contrary, as requested in our email of September 26, we have instructed the Authorized Office to go ahead with issuing an International Preliminary Examination Report. We will forward the report to you in due course.

Best regards, Paul

Paul F. Fehlner, Ph.D.
Darby & Darby P.C.
805 Third Avenue
New York, NY 10022

(212) 527-7665 | direct
(212) 753-6237 | fax

<http://www.darbylaw.com>

